(19) Eur päisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 0 960 938 A1

(12) EUROPEAN PATENT APPLICATION

published in accordance with Art. 158(3) EPC

(43) Date of publication:
01.12.1999 Bulletin 1999/48

(21) Application number: **97950408.1**

(22) Date of filing: **25.12.1997**

(51) Int. Cl.[6]: **C12N 15/54**, C12N 5/00, C12P 21/02, C12Q 1/48, C12N 9/12

(86) International application number:
PCT/JP97/04855

(87) International publication number:
WO 98/29552 (09.07.1998 Gazette 1998/27)

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(30) Priority: **27.12.1996 JP 35786496**

(71) Applicant:
Chugai Research Institute for Molecular Medicine Inc.
Niihari-gun, Ibaraki 300-41 (JP)

(72) Inventors:
- **NEZU, Jun-ichi;**
  153-2, Nagai; Niihari-mura, Niihari-gun; (JP)
- **OKU, A.;**
  Chugai Res. I. For M. Med. Inc.
  Ibaraki300-41 h (JP)

(74) Representative:
VOSSIUS & PARTNER
Siebertstrasse 4
81675 München (DE)

(54) NOVEL SERINE-THREONINE KINASE GENE

(57) A novel gene having the consensus sequence of a serine-threonine kinase active site has been isolated by the suppression subtractive hybridization method which comprises preparing a library of genes expressed specifically in fetal livers and isolating clones from this library at random. This gene presumably participates in cell growth control because it is highly expressed, especially in actively growing cells, and exhibits a significant homology with a vaccinia virus B1R kinase gene. Thus, it can be utilized as a target for developing cell growth inhibitors or antitumor agents.

EP 0 960 938 A1

Descripti n

Technical Field

[0001] The present invention relates to a novel serine-threonine kinase gene.

Background Art

[0002] Fetal tissues are comprised of many undifferentiated cells that proliferate actively, highly activated cells, nascent vascular endothelial cells, and so on. Although the activity of these cells in fetal tissues is stringently regulated and inhibited as individuals mature, the state of fetal tissues can be considered similar to that of a solid tumor except that the activity is regulated. Therefore, some of the genes expressed specifically or more strongly in fetal tissues (fetal genes) can be genes involved in the phenomena characteristic of solid tumors such as abnormal growth, immortalization, infiltration, metastasis, and angiogenesis. In addition, some diseases other than tumors are also supposed to arise because fetal genes, which are repressed in a normal living body, are abnormally activated. Therefore, genes involved in various diseases such as tumors can be screened by isolating and analyzing fetal genes.

[0003] However, there are still few reports on systematic analysis focusing merely on fetal genes from these viewpoints, and at present there is a far from perfect understanding of these gene groups.

Disclosure of the Invention

[0004] An objective of this invention is to isolate genes expressed specifically in fetal tissues and to screen genes related to diseases.

[0005] The present inventors thought that fetal tissue cells could be a model for solid tumor cells and that genes involved in diseases such as tumors could be screened by isolating and analyzing fetal genes. Furthermore, the present inventors thought it possible to develop a medicine with a novel action mechanism by designing drugs targeting the genes. Based on these thoughts, the present inventors have tried to isolate fetal genes.

[0006] Specifically, the present inventors prepared a subtraction library with many genes expressed specifically in fetal livers (or more strongly than adult livers) by the suppression subtractive hybridization method, extracted clones from this library at random, and analyzed their structure. By doing so, the present inventors succeeded in isolating a novel gene, VRK1, having the consensus sequence of a serine-threonine kinase active site. The present inventors also performed a data base search based on the amino acid sequence deduced from the isolated gene. The present inventors thus have found this gene product exhibits a significant homology with B1R kinase, which is presumably involved in DNA replication of vaccinia virus. In addition, the present inventors found human EST having a very high homology with this gene in the database and isolated its full-length cDNA, VRK2. Analyzing the expression of the two isolated genes in various cells by northern blot analysis showed that these genes are strongly expressed, especially in actively growing cells such as human fetal livers, testes, and various tumor cell lines. Furthermore, the present inventors have found that the VRK1 protein actually has protein kinase activity.

[0007] Thus, the present invention relates to novel serine-threonine kinase genus, VRK1 and VRK2. More specifically, the present invention relate to:

(1) a protein having the amino acid sequence of SEQ ID NO: 2, or a protein having the same amino acid sequence where one or more amino acids are added, deleted, or substituted and having serine-threonine kinase activity,
(2) a protein having the amino acid sequence of SEQ ID NO: 4, or a protein having the same amino acid sequence where one or more amino acids are added, deleted, or substituted and having serine-threonine kinase activity,
(3) a protein encoded by a DNA sequence that hybridizes with the DNA sequence of SEQ ID NO: 1 or its complementary sequence and having serine-threonine kinase activity,
(4) a protein encoded by a DNA sequence that hybridizes with the DNA sequence of SEQ ID NO: 3 or its complementary sequence and having serine-threonine kinase activity,
(5) a DNA encoding the protein of any one of (1) to (4),
(6) a vector comprising the DNA of (5),
(7) a transformant carrying the vector of (6),
(8) a method of producing the protein of any one of (1) to (4), wherein the method comprises cultivating the transformant of (7),
(9) an antibody binding to the protein of any one of (1) to (4),
(10) an antisense DNA against the DNA of (5) or part of it,
(11) a method of screening compounds having inhibitory activity of serine-threonine kinase activity of the protein of any one of (1) to (4), wherein the method comprises

(a) contacting the protein of any one of (1) to (4) with a substrate to be phosphorylated by this protein in the presence of a test compound to detect the kinase activity of the protein of any one of (1) to (4), and

(b) comparing the kinase activity detected in step (a) with that detected in the absence of the test compound and selecting a compound that lowers the kinase activity of the protein of any one of (1) to (4).

[0008] The present invention relates to novel serine-threonine kinases, "VRK1" and "VRK2." The nucleotide sequence of the "VRK1" cDNA and the amino acid sequence of the protein are shown in SEQ ID NO: 1 and 2, respectively. In addition, the nucleotide sequence of the "VRK2" cDNA and the amino acid sequence of the protein are shown in SEQ ID NO: 3 and 4, respectively. "VRK1" cDNA has a significant homology with B1R kinase, which is presumably involved in DNA replication of vaccinia virus. The gene is also characterized by its strong expression in actively growing cells such as fetal livers, testes, and various tumor cell lines. In addition, overexpression of "VRK1" protein drastically increases the growing activity of NIH3T3 cells. These facts imply "VRK1" is involved in the regulation mechanism of cell growth. "VRK1" protein has protein kinase activity, which presumably plays an important roll in the regulation of cell growth. "VRK2" has a high homology with "VRK1," especially in the serine-threonine kinase site. "VRK2," like "VRK1," has a significant homology with B1R kinase, and the gene is characterized by its strong expression in actively growing cells such as fetal livers, testes, and various tumor cell lines. These facts imply "VRK2" has the same function as that of "VRK1."

[0009] "VRK1" and "VRK2" proteins can be prepared as recombinant proteins with recombinant DNA techniques or as natural proteins. The recombinant proteins can be prepared, for example, by cultivating cells transformed with the DNAs encoding these proteins, as will be described later. Natural proteins can be isolated from fetal livers, testes, or tumor cell strains such as HeLa S3, in which these proteins are highly expressed, by a method well-known to one skilled in the art, such as affinity chromatography with the antibodies of the present invention as described later. Either polyclonal or monoclonal antibodies can be used. The polyclonal antibodies can be prepared from, for example, serum from small animals such as rabbits immunized with these proteins by, for example, ammonium sulfate precipitation, protein A- or protein G-column chromatography, DEAE ion exchange chromatography, affinity chromatography using a column coupled with these proteins, etc. The monoclonal antibodies can be prepared as follows. First, a small animal such as a mouse is immunized with these proteins. The spleen is extracted from the mouse and dissociated to cells. The resulting cells are fused to mouse myeloma cells using a reagent such as polyethylene glycol, and the clone that produces antibodies against these proteins is screened from the fusion cells (hybridoma) thus generated. The hybridoma thus obtained is then transplanted into a mouse abdominal cavity. Ascites is collected from the mouse and purified by, for example, ammonium sulfate precipitation, protein A- or protein G-column chromatography, DEAE ion exchange chromatography, affinity chromatography using a column coupled with "VRK1" or "VRK2" protein, etc. If the antibodies obtained are to be used for administering to a human body (for antibody therapy or the like), humanized antibodies or human antibodies should be used to decrease immunogenicity. An example of methods for humanizing antibodies is the CDR graft method, in which an antibody gene is cloned from monoclonal antibody-producing cells and its antigenic determinant is transplanted to an existing human antibody. Besides, human antibodies can be directly prepared just like usual monoclonal antibodies by immunizing a mouse whose immune system is replaced with a human immune system.

[0010] Furthermore, one skilled in the art can prepare not only natural "VRK1" and "VRK2" proteins (SEQ ID NO: 2 and 4, respectively) but also proteins with substantially the same function as that of the natural proteins, if needed, by replacing amino acids in the proteins by a well-known method. Besides, mutations of amino acids in proteins can occur naturally. Thus, mutant proteins with serine-threonine kinase activity that are generated by introducing amino acid substitution, deletion, or addition into the natural proteins are also included in the proteins of the present invention. Methods for amino acid alteration, for example, a site-directed mutagenesis system using PCR (GIBCO-BRL, Gaithersburg, Maryland), the oligonucleotide-mediated site-directed mutagenesis method (Kramer, W. and Fritz, HJ (1987) Methods in Enzymol. 154:350-367), and the Kunkel method (Methods Enzymol. 85, 2763-2766 (1988)), are well-known to one skilled in the art. Furthermore, usually ten or less, preferably six or less, and more preferably three or less amino acids are substituted. The site where substitution, deletion, or addition is introduced is not particularly limited as long as the serine-threonine kinase activity is maintained. From the viewpoint of protein activity, the addition, deletion, or substitution of amino acids should be performed in a region other than the region corresponding to the consensus sequence of a serine-threonine kinase active site and to the consensus sequence of a protein kinase ATP binding site. Moreover, serine-threonine kinase activity of a protein can be detected, for example, by the method described in Example 9, mentioned later.

[0011] Furthermore, one skilled in the art can usually isolate DNAs having a high homology with the DNA encoding "VRK1" or "VRK2" protein (SEQ ID NO: 1 or 3, respectively) based an the DNA or the part of it using a hybridization technique (Sambrook, J. et al., Molecular Cloning 2nd ed. 9.47-9.58, Cold Spring Harbor Lab. Press, 1989) and obtain proteins having substantially the same function as VRK1 or VRK2 protein (SEQ ID NO: 2 or 4, respectively) from the DNA. Thus, proteins with serine-threonine kinase activity that are encoded by DNAs hybridizing with DNA encoding "VRK1" or "VRK2" protein are also included in the proteins of the present invention. Hybridizing DNAS are isolated from

other organisms including, for example, mice, rats, rabbits, bovines, and so on. Tissues such as fetal livers and testes are especially suitable for isolating. Thus isolated DNAs encoding proteins having substantially the same function as that of "VRK1" or "VRK2" proteins usually have a high homology with the DNA (SEQ ID NO: 1 or 3) encoding "VRK1" or "VRK2" protein, respectively. The term "high homology" used herein means at least 40% or more, preferably 60% or more, and more preferably 80% or more of sequence identity at the amino acid level. From the viewpoint of the protein activity, a high homology should be found in the regions corresponding to the consensus sequence of a serine-threonine kinase active site and to the consensus sequence of a protein kinase ATP binding site.

[0012] Examples of conditions for hybridization to isolate these DNAs are as follows. After prehybridization at 55°C for 30 minutes or longer, hybridization is performed by adding labeled probes and incubating at 37 to 55°C for an hour or longer using "ExpressHyb Hybridization Solution" (CLONTECH). After that, the resulting hybridized product is washed three times for 20 minutes each at room temperature in 2 x SSC and 0.1% SDS then once at 37°C in 1 x SSC and 0.1% SDS. More preferably, after prehybridization at 60°C for 30 minutes or longer, hybridization is performed by adding labeled probes and incubating at 60°C for an hour or longer using "ExpressHyb Hybridization Solution" (CLONTECH). Thereafter, the hybridized product is washed three times for 20 minutes each at room temperature in 2 x SSC and 0.1% SDS then twice at 50°C in 1 x SSC and 0.1% SDS. Still more preferably, after prehybridization at 68°C for 30 minutes or longer, hybridization is performed by adding labeled probes and incubating at 68 °C for an hour or longer using "ExpressHyb Hybridization Solution" (CLONTECH). Thereafter, the hybridized product is washed three times for 20 minutes each at room temperature in 2 x SSC and 0.1% SDS then twice at 50°C in 0.1 x SSC and 0.1% SDS.

[0013] The present invention also relates to the DNAs encoding the above-described proteins of the present invention. The DNAs of the present invention include CDNAs, genomic DNAs, and synthetic DNAs as long as they encode the proteins of the present invention. The DNAs of the present invention can be used to produce the recombinant proteins. Specifically, the recombinant proteins can be prepared by inserting the DNA (for example, the DNA of SEQ ID NO: 1 or 3) of the present invention into a suitable expression vector, cultivating the transformant obtained by introducing the vector into suitable cells, and purifying the expressed proteins. For example, mammalian cells such as COS, CHO, or NIH3T3 cells; insect cells such as Sf9 cells; yeast cells; and *E. coli* cells can be used for producing the recombinant proteins. Vectors for expressing recombinant proteins in these cells vary depending on the host cells. For example, pcDNA3 (Invitrogen) or pEF-BOS (Nucleic Acids. Res. 1990, 18(17), p5322) is used for mammalian cells; "BAC-to-BAC baculovirus expression system" (GIBCO BRL), for insect cells; "Pichia Expression Kit" (Invitrogen), for yeast cells; and pGEX-5X-1 (Pharmacia) or "QIAexpress system" (Qiagen), for E. coli cells. Vectors can be introduced into host cells by, for example, the method using calcium phosphate, DEAE dextran, or cationic liposome DOTAP (Boehringer Manheim); electroporation; the calcium chloride method; etc. The recombinant proteins can be purified from the obtained transformants by the usual methods such as the method described in "The Qiaexpressionist handbook, Qiagen, Hilden, Germany."

[0014] Furthermore, the DNAs of the present invention can be used for gene therapy of diseases caused by mutations in genomic DNAs. In gene therapy, the DNAs of the present invention are administered to a living body inserted into adenovirus vectors (e.g., pAdexLcw), retrovirus vectors (e.g., pZIPneo) and so on. They can be administered by either *ex vivo* methods or *in vivo* methods.

[0015] Furthermore, since the proteins of the present invention are presumably involved in the regulation of cell growth, antisense DNAs against the DNAs of the present invention or part of them can be used as inhibitors for developing cell growth or as antitumor agents. The antisense DNAs are administered to a living body directly or in the form of the vectors into which they have been inserted. The antisense DNAs can be synthesized by methods well known to one skilled in the art.

[0016] The present invention also relates to a method of screening compounds having inhibitory activity of serine-threonine kinase activity of the proteins of the present invention. This screening method comprises two steps. First, the protein of the present invention is caused to contact a substrate to be phosphorylated by this protein in the presence of a test compound to detect the kinase activity of the protein of the present invention. Second, the kinase activity detected in step (a) is compared with that detected in the absence of the test compound, and a compound that lowers the kinase activity of the protein of the present invention is selected.

[0017] Test compounds used for this screening method are not particularly limited and are generally low-molecular-weight compounds, proteins (including the above-described antibodies of the present invention), peptides, etc. Test compounds are either artificially synthesized or natural. Substrates are, for example, casein, IkBα protein, etc. The kinase activity of the protein of the present invention can be detected, for example, by adding ATP having radioactively labeled phosphate to the reaction system containing the protein of the present invention and the substrate and measuring the radioactivity of the phosphate attached to the substrate. Specifically, the activity is detected by the method described in Example 9. The compounds thus isolated can be used as cell growth inhibitors or antitumor agents. Moreover, the present inventors learned that "VRK1" protein phosphorylates IkBα protein. IkBα is thought to be rapidly degraded when phosphorylated, thereby releasing and activating NF-kB bound thereto. In addition, NF-kB is well known as a central transcriptional regulator that causes widespread immune reactions and inflammation reactions.

Therefore, compounds that inhibit the kinase activity of the proteins of the present invention can be used as antiphlogistics and immunosuppressants.

Brief Description of the Drawings

[0018]

Figure 1 shows the adapters used for constructing the subtraction library.
Figure 2A shows the consensus sequence of the active site of serine-threonine kinase. Figure 2B shows the consensus sequence of the ATP binding site of protein kinase.
Figure 3 shows the nucleotide sequence of the clone fls223 and its deduced amino acid sequences.
Figure 4 shows electrophoretic patterns demonstrating the result of RT-PCR analysis performed to detect expressions of VRK1 and VRK2 genes in fetal and adult livers. In the figure, "A" and "F" represent Adult liver, and Fetal liver. "Low," "Middle," and "High" represent the level of PCR cycles.
Figure 5 compares amino acid sequences of VRK1 and B1R.
Figure 6 compares amino acid sequences of VRK1 and VRK2.
Figure 7 compares amino acid sequences of VRK2 and B1R.
Figure 8 shows electrophoretic patterns demonstrating the result of northern blot analysis of the expression of VRK1 and VRK2 genes in various cells.
Figure 9 shows an electrophoretic pattern demonstrating the result of western blot analysis using anti c-Myc antibody. Cell extracts from COS7 cells transfected with plasmid DNA, pcDNA3 (lane 1) or pcDNA3/VRK1myc (lane 2), were examined.
Figure 10 shows an electrophoretic pattern demonstrating the result of northern blot analysis using the VRK1 cDNA as a probe. Total RNA samples prepared from NIH3T3 cells transfected with plasmid DNA, pCOS (lane 1) or pCOS/VRK1w (lane 2), and from a human hepatoma cell line, HepG2 cells (lane 3), were examined.
Figure 11 presents microscopic photographs showing the result of colony assay. A pool of NIH3T3 cells transfected with plasmid DNA, pCOS ("pCOS") or pCOS/VRK1w ("pCOS/VRK1w") was examined.
Figure 12 shows an electrophoretic pattern of purified GST fusion proteins (CBB staining). Fusion proteins with wild VRK1 protein (lane 1) or with a mutant one (lane 2) were examined.
Figure 13 shows electrophoretic patterns demonstrating the result of kinase assay. Added proteins are indicated by "+" on the upper portion. Arrows indicate phosphorylated GST-VRK1 ("A," autophosphorylation), phosphorylated casein ("C"), phosphorylated GST-IkBα ("I"), and phosphorylated IkBα C-terminal peptide ("P").
Figure 14 shows an electrophoretic pattern demonstrating the result of kinase assay. Reactions were performed in the presence of various divalent cations at various concentrations as indicated on the right. Arrows indicate phosphorylated GST-VRK1 ("A," autophosphorylation), and phosphorylated casein ("C").
Figure 15 shows an electrophoretic pattern demonstrating the result of western blot analysis with an antibody against a VRK1 peptide using K562 cell extracts.

Best Mode for Implementing the Invention

[0019] The present invention is illustrated below in detail with reference to examples, but is not to be construed as being limited thereto.

Example 1

Construction of a subtraction library

[0020] A subtraction library was prepared using the PCR-Select™ cDNA Subtraction kit (CLONTECH) basically according to the method described by Luda Diatchenko et al. (Proc. Natl. Acad. Sci. USA, Vol.93, 6025-6030, 1996).

[0021] First, double-stranded cDNAs were synthesized from polyA$^+$ RNA prepared from human fetal and adult livers by the standard method using MMLV reverse transcriptase. Next, the respective CDNAs were blunt-ended with T4 DNA polymerase, then cleaved by *Rsa*I. A part of the cDNA originating from fetal liver (tester) was split in two; one of which was ligated with the adapter-1 and the other with the adapter-2 (Figure 1). Each aliquot was mixed with an excess amount of the adult liver cDNA (driver), denatured by heat, and subjected to the first hybridization at 68°C for 8 hours. Aliquots were then combined without heat denaturation, mixed further with an excess amount of heat-denatured driver, and subjected to the second hybridization at 68°C for about 16 hours. The mixture was diluted in the dilution buffer, incubated at 75°C for 7 minutes to remove the shorter strands of adapters, and used as a template for PCR. By performing PCR with primers corresponding to the adapters, "PCR primer-1" (SEQ ID NO: 5) and "PCR primer-2" (SEQ

ID NO: 6), cDNAs carrying different adapters on their two ends (subtracted cDNAS) were selectively amplified (suppression PCR). To obtain products with further selectivity, a portion of the amplified products was used as a template for PCR with primers "Nested PCR primer-1" (SEQ ID NO: 7) and "Nested PCR primer-2" (SEQ ID NO: 8), which locate further inside of the primers; "PCR primer-1" (SEQ ID NO: 5); and "PCR primer-2" (SEQ ID NO: 6). The products were purified using the "QIAquick PCR Purification kit" (QIAGEN), and cloned into the pT7Blue-T vector (Novagen) by the TA cloning method to create a subtraction library.

Example 2

Sequence analysis

[0022] Plasmid DNA prepared by the alkali SDS method or products of colony PCR were used as a template for sequence reaction. Sequence reaction was performed by the cycle-sequencing method using the ABI PRISM™ Dye Terminator Cycle Sequencing Ready Reaction Kit With AmplyTaq DNA Polymerase, FS, and the result was analyzed by the ABI 377 DNA Sequencer.

[0023] Colony PCR was performed as follows. Colonies carrying recombinant vectors were directly suspended into PCR reaction mixtures that contain vector primers, "M13 P4-22 primer" (SEQ ID NO: 9) and "M13 P5-22 primer" (SEQ ID NO: 10). After PCR reaction, amplified insert DNA was separated from unreacted primers and nucleotides by gel filtration or the like, and used as a template for sequencing.

[0024] As a result, the clone fls223 (261 bp) (later renamed "VRK1") was found to be able to encode an amino acid sequence (Figure 3) that contains the consensus sequence of the active site of serine-threonine kinase ([Leu, Ile, Val, Met, Phe, Tyr, Cys]-Xaa-[His, Tyr]-Xaa-Asp-[Leu, Ile, Val, Met, Phe, Tyr]-Lys-Xaa-Xaa-Asn-[Leu, Ile, Val, Met, Phe, Tyr, Cys, Thr]-[Leu, Ile, Val, Met, Phe, Tyr, Cys, Thr]-[Leu, Ile, Val, Met, Phe, Tyr, Cys, Thr]) (corresponds to amino acids at 173-185 of SEQ ID NO: 2) (Figure 2A). In addition, no gene registered in the database completely matches the nucleotide sequence of this clone. Thus, the gene is a novel one.

Example 3

RT-PCR assay

[0025] Using polyA$^+$ RNA extracted from fetal and adult livers, single-stranded cDNAs were synthesized by the standard method with SUPERSCRIPT™ II RNase H$^-$ Reverse Transcriptase (GIBCO BRL). Some of the cDNAs were used as a template for RT-PCR analysis of fls223. PCR was performed using TaKaRa Taq (TaKaRa) as Taq polymerase by the hot-start method, where the reaction was started by adding TaqStart™ Antibody (CLONTECH). Primers "FLS223 S1 primer" (SEQ ID NO: 11) and "FLS223 A1 primer" (SEQ ID NO: 12) were used to amplifyfls223.

[0026] The G3PDH (glyceraldehyde 3-phosphate dehydrogenase) gene, which is a housekeeping gene equally expressed in various tissues and known to be influenced only slightly by various inducers on its expression, was used as a control. G3PDH was amplified using the primers "hG3PDH5' primer" (SEQ ID NO: 13) and "hG3PDH3' primer" (SEQ ID NO: 14). RT-PCR analysis confirmed that the clone fls223 is strongly expressed in fetal liver, and its expression was also detected in adult liver (Figure 4). The full-length cDNA was then cloned for more detailed analysis of the gene.

Example 4

Cloning by Rapid Amplification of cDNA End (RACE)

[0027] The Marathon™ Ready cDNA (CLONTECH) or cDNA prepared by the Marathon™ cDNA Amplification Kit (CLONTECH) was used as a template for 5' RACE and 3' RACE (Chenchik A. et al., CLONTECHniques X, 1, 5-8, 1995).

[0028] The primers described above, "FLS223 S1 primer" (SEQ ID NO:. 11) and "FLS223 A1 primer" (SEQ ID NO: 12), were used for 5' RACE and 3' RACE of VRK1/fls223. Using a combination of these primers and a primer AP1 (SEQ ID NO: 15), corresponding to the adapter of the template cDNA, the reaction performed basically consisted of a reaction at 94°C for 2 minutes; five cycles of reactions at 94°C for 30 seconds and at 68°C 4 minutes; and 30 cycles of reactions at 94°C for 30 seconds, 62°C for 1 minute, and 72°C for 3 minutes; and a reaction at 72°C for 10 minutes. TaKaRa Ex Taq (TaKaRa) was used for PCR, and the reaction was started by the hot-start method by adding TaqStart™ Antibody (CLONTECH). After reaction, detected bands were recovered using the QIAquick Gel Extraction Kit (QIAGEN), and subcloned into the pT7Blue-T vector (Novagen).

[0029] Analysis of the entire nucleotide sequence revealed that the full-length fls223 cDNA encodes an open reading frame composed of 396 amino acids (refer to SEQ ID NO: 1). In the former part of the amino acid sequence, there exists

a consensus sequence of the ATP binding site of protein kinase ([Leu, Ile, Val]-Gly-Xaa-Gly-Xaa-[Phe, Tyr, Trp, Met, Gly, Ser, Thr, Asn, His]-[Ser, Gly, Ala]-Xaa-[Leu, Ile, Val, Cys, Ala, Thr]-Xaa-Xaa-[Gly, Ser, Thr, Ala, Cys, Leu, Ile, Val, Met, Phe, Tyr]-Xaa(five times or 18 times)-[Leu, Ile, Val, Met, Phe, Tyr, Trp, Cys, Ser, Thr, Ala, Arg]-[Ala, Ile, Val, Pro]-[Leu, Ile, Val, Met, Phe, Ala, Gly, Cys, Lys, Arg]-Lys) (corresponds to the amino acids 43-71 described in the SEQ ID NO: 2) (Figure 2B), and a consensus sequence of the active site of serine-threonine kinase, which is also found in the original clone. Thus, the gene product is assumed to be a novel serine-threonine kinase.

[0030] A homology search of the whole database revealed that the gene shows high homology to the B1R gene product of the Vaccinia virus (J. Gen. Virol., 70, 3187-3201, 1989; J. Gen. Virol., 72, 1349-1376, 1991) (Figure 5). The B1R gene encodes a protein composed of 300 amino acids and is assumed to be a serine-threonine kinase because the gene contains the consensus sequences analogous to that of the ATP binding site of protein kinase and of the active site of serine-threonine kinase. The full-length fls223 cDNA and the B1R gene showed relatively high homology over the entire region as well as in the kinase domain (Smallest Sum probability in Blast search = 2.7e-78). Therefore, the gene is named "Vaccinia virus B1R kinase related Kinase 1" (VRK1).

[0031] B1R kinase is an early gene whose expression is observed in early stages. It appears several hours after vaccinia virus infection and is then repressed. It has been shown that in a mutant strain containing a point mutation on the gene, virus replication stops during DNA replication. Thus, it has been hypothesized that B1R kinase regulates virus DNA replication (J. Biol. Chem., 264, 21458-21461, 1989).

[0032] VRK1 also exhibits an obvious homology to B1R kinase in the region outside of the serine-threonine kinase domain. Thus, VRK1 may participate in the regulation of cellular DNA replication or, more widely, in cell growth control, as is the case for B1R kinase in virus. This notion is supported by the fact that the VRK1 genes are more strongly expressed in tissues such as fetal liver and the testis, where numerous actively growing cells exist.

[0033] Furthermore, a public clone "human EST - H80169," which has an extremely high homology to VRK1, was found by searching the data base. Using the primers "RK A2 primer" (SEQ ID NO: 16) and "RK S1 primer" (SEQ ID NO: 17) for 5' RACE and 3' RACE, the full-length cDNA of the gene was cloned as described for VRK1, and the entire nucleotide sequence was determined. As a result, it was found that the gene encodes an open reading frame composed of 508 amino acids (refer to SEQ ID NO: 3), in which the consensus sequence of the active site of serine-threonine kinase exists. Thus, this gene may also encode a novel serine-threonine kinase. The amino acid sequence has an extremely high homology to VRK1, especially near the kinase domain (Figure 6), and a high homology to the vaccinia virus B1R kinase (Figure 7). These suggest a close relationship between this kinase and B1R kinase. Thus, it was named "Vaccinia virus B1R kinase related Kinase 2" (VRK2).

[0034] RT-PCR confirmed that VRK2 is also expressed more strongly in fetal liver than in adult liver (Figure 4). The primers "RK S2 primer" (SEQ ID NO: 18) and "RK A2 primer" (SEQ ID NO: 16) were used for RT-PCR.

Example 5

Chromosome mapping

[0035] Chromosome mapping of the VRK1 and VRK2 genes was performed using the GENEBRIDGE 4 Radiation Hybrid Panel (Research Genetics, Inc.) (Nature Genetics, 7, 22-28, 1994). DNA on the panel was used as a template for PCR. For VRK1, PCR was performed with a combination of the above primers ("FLS223 S1 primer" (SEQ ID NO: 9) and "FLS223 A1 primer" (SEQ ID NO: 12)) by a reaction at 94°C for 5 minutes; five cycles of reactions at 94°C for 30 seconds and at 72°C for 2 minutes; and 30 cycles of reactions at 94°C for 30 seconds and at 68°C for 2 minutes; and a reaction at 72°C for 3 minutes. For VRK2, PCR was performed with a combination of primers "VRK2 A primer" (SEQ ID NO: 19) and "VRK2 B primer" (SEQ ID NO: 20) by a reaction at 94°C for 3 minutes; 30 cycles of reactions at 94°C for 30 seconds, 60°C for 1 minute, and 72°C for 2 minutes; and a reaction at 72°C for 5 minutes. The resulting pattern was analyzed on a database on Internet (http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl), and maps were obtained.

[0036] The VRK1 gene was thus mapped to the position between the STS markers "D14S265" and "APM063XE7" on chromosome 14. Similarly, the VRK2 gene was mapped to the position between the STS markers "CHLC.GATA23H01" and "D2S357" on chromosome 2.

Example 6

Northern blot analysis

[0037] The expressions of VRK1 and VRK2 mRNA in various human normal tissues and tumor cell lines were analyzed by northern blotting (Figure 8).

[0038] The 5' terminal fragment of the VRK1 cDNA (upstream region of the HindIII site at nucleotide residue 546) or

that of the VRK2 cDNA (upstream region of the *Eco*RI site at position 426) was labeled with [α-^{32}P]dCTP by the random primer method using Ready-to-Go DNA labeling beads (Pharmacia), and used as a probe. Hybridization was performed at 68°C in ExpressHyb Hybridization Solution (CLONTECH) using Multiple Tissue Northern (MTN) Blot - Human, Human II, Human Fetal II, and Human Cell line (CLONTECH) according to the method recommended by the manufacturer. Final wash was done at 50°C in 0.1 x SSC, 0.1% SDS, and the image on the filter processed through hybridization was analyzed with the BAS-2000II bioimaging analyzer (Fuji Photo Film).

[0039] In Figure 8, tumor cell lines used are malignant melanoma cells "G361," lung carcinoma cells "A549," colorectal adenocarcinoma cells "SW480," Burkitt's lymphoma cells "Raji," acute lymphoblastic leukemia cells (T cell) "MOLT-4," chronic myelogenous leukemia cells "K-562," uterocervical carcinoma cells "HeLaS3," and promyelocytic leukemia cells "HL60".

[0040] The results revealed that VRK1 was expressed relatively highly in fetal tissues, and extremely highly in fetal liver. VRK1 was expressed weakly in almost all adult tissues, but it was expressed strongly in the testis and the thymus. In tumor cell lines, a very strong expression of VRK1 was observed in six out of eight cell lines.

[0041] The expression pattern at VRK2 was basically similar to that of VRK1; VRK2 was expressed strongly in fetal liver and the testis. Similarly, it was expressed strongly in tumor cell lines. However, VRK2 was not expressed in MOLT-4 cells, in contrast to the pattern of VRK1 mRNA.

Example 7

Constructing expression plasmid DNAs

[0042] cDNA containing the entire coding region of VRK1 or VRK2 was amplified by PCR using a combination of primers, VRK1 S1 primer (SEQ ID NO: 21) and VRK1 A1 primer (SEQ ID NO: 22), or VRK2 S1 primer (SEQ ID NO: 23) and VRK2 A1 primer (SEQ ID NO: 24) from cDNA synthesized from polyA$^+$ RNA extracted from human fetal liver. The amplified product was cleaved at the *Not*I site that is attached to the end of the primers and purified by agarose gel electrophoresis to obtain DNA fragments of the correct size. These were subcloned into the pCOS vector, which was pretreated with *Not*I and dephosphorylated on its ends with alkaline-phosphatase/CIAP (TaKaRa). This vector contains an EF1α promoter and enables expressing cloned cDNA strongly in a broad range of mammalian cell lines. By sequencing the thus-obtained subclones, clones (pCOS/VRK1w, pCOS/VRK2w) without mutation such as PCR error were selected and used for expression as described below and for further construction of expression plasmid DNA.

[0043] Plasmids for expressing proteins in which the anti c-Myc antibody epitope sequence (SEQ ID NO: 25) is attached to the C-terminus were constructed as follows. Using about 50 nanograms of the plasmid DNA and with pCOS/VRK1w or pCOS/VRK2w as a temptate, PCR was performed with a combination of primers. These included VRK1 MYC1 primer (SEQ ID NO: 26) and VRK1 MYC2 primer (SEQ ID NO: 27), or VRK2 MYC1 primer (SEQ ID NO: 28) and VRK2 MYC2 primer (SEQ ID NO: 29), and cDNA with the anti c-Myc antibody epitope attached to the C-terminus of the coding sequence was amplified. KOD DNA polymerase (TOYOBO) was used as the DNA polymerase. The amplified product was cleaved at the *Bam*HI site that is attached to the end of the primers and purified by agarose gel electrophoresis to obtain DNA fragments of the correct size. These were subcloned into the pcDNAS vector (Invitrogen), which was digested with *Bam*HI and dephosphorylated on its ends with alkaline-phosphatase/CIAP (TaKaRa). By sequencing the thus-obtained subclones, clones (pcDNA3/VRK1myc, pcDNA3/VRK2myc) without mutation such as PCR error were selected, and used for later experiments.

[0044] Expression plasmid DNAs for glutathione-S-transferase (GST) fusion proteins in *E. coli* were constructed as follows. Using the plasmid DNA pCOS/VRK1w or pCOS/VRK2w as a template, the coding region was amplified by PCR with a combination of primers, VRK1 H3 primer (SEQ ID NO: 30) and VRK1 H4 primer (SEQ ID NO: 31), or VRK2 H3 primer (SEQ ID NO: 32) and VRK2 H4 primer (SEQ ID NO: 33). The amplified product was cleaved at the *Bam*HI site that is attached to the end of the primers purified by agarose gel electrophoresis to obtain DNA fragments of the correct size. These were then subcloned into the pGEX-5X-1 vector (Pharmacia), which was digested with *Bam*HI and dephosphorylated on its ends with alkaline-phosphatase/CIAP (TaKaRa). By sequencing the thus-obtained subclones, clones (pGEX/VRK1w, pGEX/VRK2w) without mutation such as PCR error were selected and used for later experiments.

[0045] A clone with a mutation introduced to the predicted ATP binding site within the kinase catalytic domain (Lys at position 71 in the amino acid sequence of SEQ ID NO: 2 is replaced by Trp) was constructed using the Chameleon™ Double-Stranded Site-Directed Mutagenesis Kit (STRATAGENE) as follows. About one microgram of the pGEX/VRK1w plasmid DNA was mixed with primers VRK1 KW1 primer (SEQ ID NO: 34) and a selection primer, Select1 primer (SEQ ID NO: 35), and heat denatured by boiling for 5 minutes. Plasmid DNA and both primers containing mutation were then annealed by incubating at room temperature for 30 minutes. Next, new DNA strands were synthesized from primers by adding substrate nucleotides, DNA polymerase, etc. These were treated with *Pst*I to digest wild plasmid DNA, and introduced into XLmutS competent cells. After overnight liquid culture, plasmid DNA was extracted then treated with PstI to digest contaminating wild plasmid DNA. It was then reintroduced into the competent cells. By sequencing sev-

eral single isolated colonies, a clone (pGEX/VRK1K71W) with an introduced mutation was selected.

Example 8

Expression in mammalian cell lines

[0046] About 10 micrograms of plasmid DNA, pcDNA3/VRK1myc or pcDNA3, was introduced (transfected) into COS7 cells using SuperFect (QIAGEN). Specifically, about 10^6 COS7 cells were plated in a 10-cm dish, and cultured overnight. A mixture of 10 micrograms of plasmid DNA and 60 microliters of SuperFect was then added to the culture, and the culture was incubated for about 3 hours. Thereafter, the culture medium was replaced with fresh medium. The cells were then cultured for two more days and collected by detaching in a trypsin-EDTA solution. Cells were washed once in PBS, disrupted in RIPA buffer (1% NP-40, 10 mM Tris-HCl (pH 7.2), 0.1% sodium deoxychorate, 0.1% SDS, 0.15 M sodium chloride, 1 mM EDTA, 10 micrograms/ml aprotinin, 1 mM PMSF), and cell extracts were obtained by centrifugation. The cell extracts were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and subjected to western blot analysis using anti c-Myc antibody (SANTA CRUZ). A band of about 50 kDa was specifically observed in cells transfected with the pcDNA3/VRK1myc plasmid DNA, indicating that VRK1myc protein is expressed (Figure 9).

[0047] Next, about 7.5 micrograms of plasmid DNA, pCOS/VRK1w or pCOS, was transfected into NIH3T3 cells by the method using cationic phospholipid DOTAP (Boehringer Mannheim). After transfection, transformants were selected by adding G418 (GIBCO-BRL) to the culture medium to a final concentration of 500 micrograms/ml. Total RNA was prepared from each pool of transformants by the method using ISOGEN (Wajo Junyaku). The total RNA was then subjected to northern blot analysis using the VRK1 cDNA as a probe. The results confirmed that VRK1 mRNA was expressed in a pool of cells obtained by transfection with the pCOS/VRK1w plasmid DNA (Figure 10). These pools of cells were examined for the ability to form colonies on soft agar (colony assay). To this end, 2×10^4 cells were suspended in Dulbecco's Modified Eagle Medium (DMEM) containing 10% fetal bovine serum and 0.4% thawed SeaPlaque agarose (TaKaRa), and overlaid on a bottom agarose which was made of 0.53% SeaPlaque agarose, 10% fetal bovine serum, and DMEM. After two-week culturing, the pool of cells obtained by transfection with the pCOS/VRK1w plasmid DNA formed a number of colonies that were obviously larger than those formed in the pool of cells obtained by transfection with the pCOS plasmid DNA. This suggests that overexpression of VRK1 confers abnormal growth activities on cells (Figure 11).

Example 9

Expression of VRK1 protein in *E.coli* and kinase assay

[0048] Both wild VRK1 protein and mutant VRK1 protein were expressed in *E.coli* as a fusion protein with GST protein and purified. The *E.coli* DH5α strain cells carrying the above-described plasmid DNA, pGEX/VRK1W, or pGEX/VRK1K71W were cultured overnight at 37°C in 10 ml 2xYT medium. Some of the culture was diluted 100-fold with fresh 2xYT medium and cultured at 37°C until the OD value at 600 nm reached 0.6. IPTG (isopropyl-β-D(-)-thiogalactopyranoside) was then added to the culture to a final concentration of 0.1 mM, and the culture was incubated further for several hours. The *E. coli* cells were collected by centrifugation, resuspended in PBS containing 1% Triton X-100 and 1% Tween 20, and subsequently disrupted by sonication to solubilize proteins. From the solubilized samples, wild VRK1 protein and mutant VRK1, which were expressed as a fusion protein with GST, were purified by affinity chromatography using glutathione Sepharose4B (Pharmacia). These proteins were subjected to SDS-PAGE and stained with Coomassie Brilliant Blue (CBB) to confirm their purity (Figure 12). GST protein and GST-IkBα protein were prepared in the same manner.

[0049] Kinase assay was performed on a total of 50 microliters of a reaction mixture containing 0.2 micrograms of wild or mutant VRK1 protein, 50 mM Tris-HCl (pH 7.2), 1 mM dithiothreitol (DTT), 2 mM or 10 mM of divalent cation (Mg, Mn, Zn, Ca), a maximum of 5 micrograms of substrate protein, and 1 microliter of [γ-^{32}P]-ATP (3000 Ci/mM, 10 mCi/ml [Amersham]). In some experiments, another buffer system containing 40 mM Hepes (pH 7.4), 1 mM DTT, and 2.5 mM EGTA was used.

[0050] Specifically, the reaction was first performed in the presence of 10 mM Mg at 37°C for 30 minutes using, as a protein substrate, histone (Nakalai), casein (Sigma), myelin basic protein/MBP (Sigma), GST, GST-IkBα, or IkBα C-terminal peptide (SEQ ID NO: 36). The reaction mixture was subjected to SDS-PAGE, and the radioactivity of phosphorylated proteins was analyzed with a BAS2000II bioimaging analyzer (Fuji Photo Film). The result indicates that wild VRK1 protein phosphorylates casein and GST-IkBα (Figure 13). In contrast, no phosphorylation was observed in reactions with mutant VRK1 carrying mutation in the predicted ATP binding site. This indicates that VRK1 is a protein kinase that contains a typical catalytic domain. In addition, GST protein was not phosphorylated by VRK1, suggesting that the phosphorylation of GST-IkBα protein by VRK1 occurs within IkBα protein but not in GST moiety.

[0051] IkBα is said to negatively regulate the function of transcription factor NF-kB by forming a complex with it. In addition, it is widely accepted that IkBα is inactivated by self-phosphorylation, immediately thereafter undergoes proteolysis, thereby liberating and activating NF-kB. NF-kB is supposed to be a central transcriptional regulator that induces a broad range of immune reactians and inflammatory reactions. Therefore, a kinase that phosphorylates IkBα is important as a target molecule of anti-inflammatory drugs. Since VRK1 strongly phosphorylates IkBα *in vitro*, VRK1 probably participates in the activation of NF-kB by phosphorylating IkBα *in vivo* as well. Therefore, it is possible to anticipate anti-inflammatory effects or immunosuppressive effects by inhibiting VRK1 kinase activity, or by reducing its protein amount.

[0052] Next, the requirement for divalent cations in phosphorylation by VRK1 was examined (Figure 14). In the presence of various divalent cations (Mg, Mn, Zn, and Ca) at a final concentration of 2 mM or 10 mM, kinase reactions were performed using casein as a substrate protein. The result showed that VRK1 exhibits phosphorylation activity in the presence of all divalent cations except for Zn. However, the levels of activity were different; VRK1 exhibited especially strong activity in the presence of Mn.

Example 10

Preparation of an antibody against VRK1 protein

[0053] A peptide (SEQ ID NO: 37) corresponding to the C-terminal sequence of the deduced VRK1 amino acid sequence was synthesized (Sawady Technology) and conjugated to Keyhole limpet hemocyanin (KLH) at its amino-terminal cysteine mediated by m-maleimidobenzoyl-N-hydroxy-succinimide ester (MBS). This was used as an antigen to immunize rabbits, and antisera were obtained. Antibodies that specifically react with the peptide were purified from the antisera by affinity chromatography using Cellulofine (Seikagaku Corporation) conjugated with the peptide. Western blot analysis using extracts of K562 cells, which were confirmed by northern blot analysis to strongly express VRK1, detected a single band with a molecular weight of approximately 50 kDa, indicating that VRK1 protein is specifically recognized by the antibody (Figure 15).

Industrial Applicability

[0054] The serine-threonine kinase genes isolated by the present inventors show significant homology to a vaccinia virus gene that is thought to be involved in DNA replication and are strongly expressed in actively growing cells. Furthermore, overexpression of proteins encoded by the genes remarkably enhances cell proliferation activity. Thus, the isolated serine-threonine kinase genes are assumed to participate in the regulation of cell growth. Therefore, it is possible to develop cell growth inhibitors or antitumor agents based on a novel mechanism by screening drugs targeted on the genes (such as antisense DNA), or drugs which can regulate either the expression of the genes or the activity of the proteins encoded by the genes of the present invention.

Sequence Listing

(1) Name of Applicant: Chugai Research Institute for Molecular Medicine, Inc.

(2) Title of the Invention: Novel Serine-threonine Kinase Gene

(3) Reference Number: C1-805PCT

(4) Application Number:

(5) Filing Date:

(6) Country where the priority application was filed and the application number of the priority application: Japan, No. Hei 8-357864

(7) Priority date: December 26, 1996

(8) Number of Sequences: 37

SEQ ID NO: 1

SEQUENCE LENGTH: 1662

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE:

NAME/KEY: CDS

LOCATION: 76..1266

SEQUENCE DESCRIPTION:

CCGAGTTACG AGTCGGCGAA AGCGGCGGGA AGTTCGTACT GGGCAGAACG CGACGGGTCT 60

GCGGCTTAGG TGAAA ATG CCT CGT GTA AAA GCA GCT CAA GCT GGA AGA CAG 111

Met Pro Arg Val Lys Ala Ala Gln Ala Gly Arg Gln

```
                1               5                  10
AGC TCT GCA AAG AGA CAT CTT GCA GAA CAA TTT GCA GTT GGA GAG ATA      159
Ser Ser Ala Lys Arg His Leu Ala Glu Gln Phe Ala Val Gly Glu Ile
        15                  20                  25
ATA ACT GAC ATG GCA AAA AAG GAA TGG AAA GTA GGA TTA CCC ATT GGC      207
Ile Thr Asp Met Ala Lys Lys Glu Trp Lys Val Gly Leu Pro Ile Gly
    30                  35                  40
CAA GGA GGC TTT GGC TGT ATA TAT CTT GCT GAT ATG AAT TCT TCA GAG      255
Gln Gly Gly Phe Gly Cys Ile Tyr Leu Ala Asp Met Asn Ser Ser Glu
45                  50                  55                  60
TCA GTT GGC AGT GAT GCA CCT TGT GTT GTA AAA GTG GAA CCC AGT GAC      303
Ser Val Gly Ser Asp Ala Pro Cys Val Val Lys Val Glu Pro Ser Asp
                65                  70                  75
AAT GGA CCT CTT TTT ACT GAA TTA AAG TTC TAC CAA CGA GCT GCA AAA      351
Asn Gly Pro Leu Phe Thr Glu Leu Lys Phe Tyr Gln Arg Ala Ala Lys
            80                  85                  90
CCA GAG CAA ATT CAG AAA TGG ATT CGT ACC CGT AAG CTG AAG TAC CTG      399
Pro Glu Gln Ile Gln Lys Trp Ile Arg Thr Arg Lys Leu Lys Tyr Leu
        95                  100                 105
GGT GTT CCT AAG TAT TGG GGG TCT GGT CTA CAT GAC AAA AAT CGA AAA      447
Gly Val Pro Lys Tyr Trp Gly Ser Gly Leu His Asp Lys Asn Gly Lys
    110                 115                 120
AGT TAC AGG TTT ATG ATA ATG GAT CGC TTT GGG AGT GAC CTT CAG AAA      495
Ser Tyr Arg Phe Met Ile Met Asp Arg Phe Gly Ser Asp Leu Gln Lys
125                 130                 135                 140
```

```
ATA TAT GAA GCA AAT GCC AAA AGG TTT TCT CGG AAA ACT GTC TTG CAG      543
Ile Tyr Glu Ala Asn Ala Lys Arg Phe Ser Arg Lys Thr Val Leu Gln
                145                 150                 155

CTA AGC TTA AGA ATT CTG GAT ATT CTG GAA TAT ATT CAC GAG CAT GAG      591
Leu Ser Leu Arg Ile Leu Asp Ile Leu Glu Tyr Ile His Glu His Glu
            160                 165                 170

TAT GTG CAT GGA GAT ATC AAG GCC TCA AAT CTT CTT CTG AAC TAC AAG      639
Tyr Val His Gly Asp Ile Lys Ala Ser Asn Leu Leu Leu Asn Tyr Lys
        175                 180                 185

AAT CCT GAC CAG GTG TAC TTG GTA GAT TAT GGC CTT GCT TAT CGG TAC      687
Asn Pro Asp Gln Val Tyr Leu Val Asp Tyr Gly Leu Ala Tyr Arg Tyr
    190                 195                 200

TGC CCA GAA GGA GTT CAT AAA GAA TAC AAA GAA GAC CCC AAA AGA TGT      735
Cys Pro Glu Gly Val His Lys Glu Tyr Lys Glu Asp Pro Lys Arg Cys
205                 210                 215                 220

CAC GAT GGC ACT ATT GAA TTC ACG AGC ATC GAT GCA CAC AAT GGT GTG      783
His Asp Gly Thr Ile Glu Phe Thr Ser Ile Asp Ala His Asn Gly Val
                225                 230                 235

GCC CCA TCA AGA CGT GGT GAT TTG GAA ATA CTT GGT TAT TGC ATG ATC      831
Ala Pro Ser Arg Arg Gly Asp Leu Glu Ile Leu Gly Tyr Cys Met Ile
            240                 245                 250

CAA TGG CTT ACT GGC CAT CTT CCT TGG GAG GAT AAT TTG AAA GAT CCT      879
Gln Trp Leu Thr Gly His Leu Pro Trp Glu Asp Asn Leu Lys Asp Pro
        255                 260                 265

AAA TAT GTT AGA GAT TCC AAA ATT AGA TAC AGA GAA AAT ATT GCA AGT      927
```

```
Lys Tyr Val Arg Asp Ser Lys Ile Arg Tyr Arg Glu Asn Ile Ala Ser
    270                 275                 280

TTG ATG GAC AAA TGT TTT CCT GAG AAA AAC AAA CCA GGT GAA ATT GCC    975
Leu Met Asp Lys Cys Phe Pro Glu Lys Asn Lys Pro Gly Glu Ile Ala
285                 290                 295                 300

AAA TAC ATG GAA ACA GTG AAA TTA CTA GAC TAC ACT GAA AAA CCT CTT   1023
Lys Tyr Met Glu Thr Val Lys Leu Leu Asp Tyr Thr Glu Lys Pro Leu
                305                 310                 315

TAT GAA AAT TTA CGT GAC ATT CTT TTG CAA GGA CTA AAA GCT ATA GGA   1071
Tyr Glu Asn Leu Arg Asp Ile Leu Leu Gln Gly Leu Lys Ala Ile Gly
            320                 325                 330

AGT AAG GAT GAT GGC AAA TTG GAC CTC AGT GTT GTG GAG AAT GGA GGT   1119
Ser Lys Asp Asp Gly Lys Leu Asp Leu Ser Val Val Glu Asn Gly Gly
        335                 340                 345

TTG AAA GCA AAA ACA ATA ACA AAG AAG CGA AAG AAA GAA ATT GAA GAA   1167
Leu Lys Ala Lys Thr Ile Thr Lys Lys Arg Lys Lys Glu Ile Glu Glu
    350                 355                 360

AGC AAG GAA CCT GGT GTT GAA GAT ACG GAA TGG TCA AAC ACA CAG ACA   1215
Ser Lys Glu Pro Gly Val Glu Asp Thr Glu Trp Ser Asn Thr Gln Thr
365                 370                 375                 380

GAG GAG GCC ATA CAG ACC CGT TCA AGA ACC AGA AAG AGA GTC CAG AAG   1263
Glu Glu Ala Ile Gln Thr Arg Ser Arg Thr Arg Lys Arg Val Gln Lys
                385                 390                 395

TAATTCAGAT GCTGTGAACC AGATTTCCTT TTCTTTGTTT TCTTTTGACT TTTTTCTCCT  1323

TTTCTGTTAG AACTGTTTTA TTTTCCTGTG AGTCTTGCGA GGTGGAATTA ATGATTAAAT  1383
```

ACTCATGTGT TCAGAAAACA TAAACTTTTT TTATAAAAAT ATTTTGTACA ATTCATTAAA 1443

GGCTAATTTA TGAAATTTGA AAATCTTCAG GTTATACTCC TTAAGTTATC CCAAAGCCGT 1503

GTGTTTGTGA TGTTTTGGAG TACATATATA TGAAAATTAT TATGACACGC ACTTTTCTAA 1563

TCATTGTACA TTTCTCAGAG TGGATAAAAA TGTTTGACAA AGTCCTCACT TTTAAGGAAA 1623

TGCAAAGCTT AAAATAAAAC TCTCTTTTGT TTGATGCAG 1662

SEQ ID NO: 2

SEQUENCE LENGTH: 396

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Met Pro Arg Val Lys Ala Ala Gln Ala Gly Arg Gln
1 5 10

Ser Ser Ala Lys Arg His Leu Ala Glu Gln Phe Ala Val Gly Glu Ile
15 20 25

Ile Thr Asp Met Ala Lys Lys Glu Trp Lys Val Gly Leu Pro Ile Gly
30 35 40

Gln Gly Gly Phe Gly Cys Ile Tyr Leu Ala Asp Met Asn Ser Ser Glu
45 50 55 60

Ser Val Gly Ser Asp Ala Pro Cys Val Val Lys Val Glu Pro Ser Asp
65 70 75

Asn Gly Pro Leu Phe Thr Glu Leu Lys Phe Tyr Gln Arg Ala Ala Lys
80 85 90

```
Pro Glu Gln Il  Gln Lys Trp Il  Arg Thr Arg Lys Leu Lys Tyr Leu
        95                  100                 105
Gly Val Pro Lys Tyr Trp Gly Ser Gly Leu His Asp Lys Asn Gly Lys
    110                 115                 120
Ser Tyr Arg Phe Met Ile Met Asp Arg Phe Gly Ser Asp Leu Gln Lys
125                 130                 135                 140
Ile Tyr Glu Ala Asn Ala Lys Arg Phe Ser Arg Lys Thr Val Leu Gln
                145                 150                 155
Leu Ser Leu Arg Ile Leu Asp Ile Leu Glu Tyr Ile His Glu His Glu
            160                 165                 170
Tyr Val His Gly Asp Ile Lys Ala Ser Asn Leu Leu Leu Asn Tyr Lys
        175                 180                 185
Asn Pro Asp Gln Val Tyr Leu Val Asp Tyr Gly Leu Ala Tyr Arg Tyr
    190                 195                 200
Cys Pro Glu Gly Val His Lys Glu Tyr Lys Glu Asp Pro Lys Arg Cys
205                 210                 215                 220
His Asp Gly Thr Ile Glu Phe Thr Ser Ile Asp Ala His Asn Gly Val
                225                 230                 235

Ala Pro Ser Arg Arg Gly Asp Leu Glu Ile Leu Gly Tyr Cys Met Ile
            240                 245                 250
Gln Trp Leu Thr Gly His Leu Pro Trp Glu Asp Asn Leu Lys Asp Pro
        255                 260                 265
Lys Tyr Val Arg Asp Ser Lys Ile Arg Tyr Arg Glu Asn Ile Ala Ser
    270                 275                 280
```

Leu Met Asp Lys Cys Phe Pro Glu Lys Asn Lys Pro Gly Glu Ile Ala
285 290 295 300
Lys Tyr Met Glu Thr Val Lys Leu Leu Asp Tyr Thr Glu Lys Pro Leu
305 310 315
Tyr Glu Asn Leu Arg Asp Ile Leu Leu Gln Gly Leu Lys Ala Ile Gly
320 325 330
Ser Lys Asp Asp Gly Lys Leu Asp Leu Ser Val Val Glu Asn Gly Gly
335 340 345
Leu Lys Ala Lys Thr Ile Thr Lys Lys Arg Lys Lys Glu Ile Glu Glu
350 355 360
Ser Lys Glu Pro Gly Val Glu Asp Thr Glu Trp Ser Asn Thr Gln Thr
365 370 375 380
Glu Glu Ala Ile Gln Thr Arg Ser Arg Thr Arg Lys Arg Val Gln Lys
385 390 395

SEQ ID NO: 3

SEQUENCE LENGTH: 1833

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

FEATURE:

NAME/KEY: CDS

LOCATION: 131..1657

SEQUENCE DESCRIPTION:

```
CTGCACTGCG AGGCCGACGC AGCTGGAGAG AAGTTAGGCA GGTCCTAGGG AGGGCAGGCT     60

CGAGTGCTGG GCCCGCCTCC CCGCGGGACT GTAGGCCCGG GGGCTCCGCC TCGTCGCAGC    120

GGCAGAAGTG ATG CCA CCA AAA AGA AAT GAA AAA TAC AAA CTT CCT ATT        169

          Met Pro Pro Lys Arg Asn Glu Lys Tyr Lys Leu Pro Ile

          1               5                  10

CCA TTT CCA GAA GGC AAG GTT CTG GAT GAT ATG GAA GGC AAT CAG TGG       217

Pro Phe Pro Glu Gly Lys Val Leu Asp Asp Met Glu Gly Asn Gln Trp

     15                  20                  25

GTA CTG GGC AAG AAG ATT GGC TCT GGA GGA TTT GGA TTG ATA TAT TTA       265

Val Leu Gly Lys Lys Ile Gly Ser Gly Gly Phe Gly Leu Ile Tyr Leu

30                  35                  40                  45

GCT TTC CCC ACA AAT AAA CCA GAG AAA GAT GCA AGA CAT GTA GTA AAA       313

Ala Phe Pro Thr Asn Lys Pro Glu Lys Asp Ala Arg His Val Val Lys

                50                  55                  60

GTC GAA TAT CAA GAA AAT GGC CCG TTA TTT TCA GAA CTT AAA TTT TAT       361

Val Glu Tyr Gln Glu Asn Gly Pro Leu Phe Ser Glu Leu Lys Phe Tyr

            65                  70                  75

CAG AGA GTT GCA AAA AAA GAC TGT ATC AAA AAG TGG ATA GAA CGC AAA       409

Gln Arg Val Ala Lys Lys Asp Cys Ile Lys Lys Trp Ile Glu Arg Lys

        80                  85                  90

CAA CTT GAT TAT TTA GGA ATT CCT CTG TTT TAT GGA TCT GGT CTG ACT       457

Gln Leu Asp Tyr Leu Gly Ile Pro Leu Phe Tyr Gly Ser Gly Leu Thr

     95                  100                 105

GAA TTC AAG GGA AGA AGT TAC AGA TTT ATG GTA ATG GAA AGA CTA GGA       505

Glu Phe Lys Gly Arg Ser Tyr Arg Phe Met Val Met Glu Arg Leu Gly
```

```
                110                 115                 120                 125
ATA GAT TTA CAG AAG ATC TCA GGC CAG AAT GGT ACC TTT AAA AAG TCA      553
Ile Asp Leu Gln Lys Ile Ser Gly Gln Asn Gly Thr Phe Lys Lys Ser
                    130                 135                 140
ACT GTC CTG CAA TTA GGT ATC CGA ATG TTG GAT GTA CTG GAA TAT ATA      601
Thr Val Leu Gln Leu Gly Ile Arg Met Leu Asp Val Leu Glu Tyr Ile
            145                 150                 155
CAT GAA AAT GAA TAT GTT CAT GGT GAT GTA AAA GCA GCA AAT CTA CTT      649
His Glu Asn Glu Tyr Val His Gly Asp Val Lys Ala Ala Asn Leu Leu
        160                 165                 170
TTG GGT TAC AAA AAT CCA GAC CAG GTT TAT CTT GCA GAT TAT GGA CTT      697
Leu Gly Tyr Lys Asn Pro Asp Gln Val Tyr Leu Ala Asp Tyr Gly Leu
    175                 180                 185
TCC TAC AGA TAT TGT CCC AAT GGG AAC CAC AAA CAG TAT CAG GAA AAT      745
Ser Tyr Arg Tyr Cys Pro Asn Gly Asn His Lys Gln Tyr Gln Glu Asn
190                 195                 200                 205
CCT AGA AAA GGC CAT AAT GGG ACA ATA GAG TTT ACC AGC TTG GAT GCC      793
Pro Arg Lys Gly His Asn Gly Thr Ile Glu Phe Thr Ser Leu Asp Ala
                210                 215                 220
CAC AAG GGA GTA GCC TTG TCC AGA CGA AGT GAC GTT GAG ATC CTC GGC      841
His Lys Gly Val Ala Leu Ser Arg Arg Ser Asp Val Glu Ile Leu Gly
            225                 230                 235
TAC TGC ATG CTG CGG TGG TTG TGT GGC AAA CTT CCC TGG GAA CAG AAC      889
Tyr Cys Met Leu Arg Trp Leu Cys Gly Lys Leu Pro Trp Glu Gln Asn
        240                 245                 250
```

```
CTG AAG GAC CCT GTG GCT GTG CAG ACT GCT AAA ACA AAT CTG TTG GAC      937
Leu Lys Asp Pro Val Ala Val Gln Thr Ala Lys Thr Asn Leu Leu Asp
    255                 260                 265

GAG CTC CCC CAG TCA GTG CTT AAA TGG GCT CCT TCT GGA AGC AGT TGC      985
Glu Leu Pro Gln Ser Val Leu Lys Trp Ala Pro Ser Gly Ser Ser Cys
270                 275                 280                 285

TGT GAA ATA GCC CAA TTT TTG GTA TGT GCT CAT AGT TTA GCA TAT GAT     1033
Cys Glu Ile Ala Gln Phe Leu Val Cys Ala His Ser Leu Ala Tyr Asp
                290                 295                 300

GAA AAG CCA AAC TAT CAA GCC CTC AAG AAA ATT TTG AAC CCT CAT GGA     1081
Glu Lys Pro Asn Tyr Gln Ala Leu Lys Lys Ile Leu Asn Pro His Gly
            305                 310                 315

ATA CCT TTA GGA CCA CTG GAC TTT TCC ACA AAA GGA CAG AGT ATA AAT     1129
Ile Pro Leu Gly Pro Leu Asp Phe Ser Thr Lys Gly Gln Ser Ile Asn
        320                 325                 330

GTC CAT ACT CCA AAC AGT CAA AAA GTT GAT TCA CAA AAG GCT GCA ACA     1177
Val His Thr Pro Asn Ser Gln Lys Val Asp Ser Gln Lys Ala Ala Thr
    335                 340                 345

AAG CAA GTC AAC AAG GCA CAC AAT AGG TTA ATC GAA AAA AAA GTC CAC     1225
Lys Gln Val Asn Lys Ala His Asn Arg Leu Ile Glu Lys Lys Val His
350                 355                 360                 365

AGT GAG AGA AGC GCT GAG TCC TGT GCA ACA TGG AAA GTG CAG AAA GAC     1273
Ser Glu Arg Ser Ala Glu Ser Cys Ala Thr Trp Lys Val Gln Lys Glu
                370                 375                 380

GAG AAA CTG ATT GGA TTG ATG AAC AAT GAA GCA GCT CAG GAA AGC ACA     1321
```

```
Glu Lys Leu Ile Gly Leu Met Asn Asn Glu Ala Ala Gln Glu Ser Thr
                385                 390                 395

AGG AGA AGA CAG AAA TAT CAA GAG TCT CAA GAA CCT TTG AAT GAA GTA    1369
Arg Arg Arg Gln Lys Tyr Gln Glu Ser Gln Glu Pro Leu Asn Glu Val
        400                 405                 410

AAC AGT TTC CCA CAA AAA ATC AGC TAT ACA CAA TTC CCA AAC TCA TTT    1417
Asn Ser Phe Pro Gln Lys Ile Ser Tyr Thr Gln Phe Pro Asn Ser Phe
    415                 420                 425

TAT GAG CCT CAT CAA GAT TTT ACC AGT CCA GAT ATA TTC AAG AAG TCA    1465
Tyr Glu Pro His Gln Asp Phe Thr Ser Pro Asp Ile Phe Lys Lys Ser
430                 435                 440                 445

AGA TCT CCA TCT TGG TAT AAA TAC ACT TCC ACA GTC AGC ACG GGG ATC    1513
Arg Ser Pro Ser Trp Tyr Lys Tyr Thr Ser Thr Val Ser Thr Gly Ile
                450                 455                 460

ACA GAC TTA GAA AGT TCA ACT GGA CTT TGG CCT ACA ATT TCC CAG TTT    1561
Thr Asp Leu Glu Ser Ser Thr Gly Leu Trp Pro Thr Ile Ser Gln Phe
            465                 470                 475

ACT CTT AGT GAA GAG ACA AAC GCA GAT GTT TAT TAT TAT CGC ATC ATC    1609
Thr Leu Ser Glu Glu Thr Asn Ala Asp Val Tyr Tyr Tyr Arg Ile Ile
        480                 485                 490

ATA CCT GTC CTT TTG ATG TTA GTA TTT CTT GCT TTA TTT TTT CTC        1654
Ile Pro Val Leu Leu Met Leu Val Phe Leu Ala Leu Phe Phe Leu
    495                 500                 505

TGAAGATGAT ACCAAAATTC CTTTTGATAA TTTTTTAAGT TTCCAGCTCT TCACCGAAAT   1714

GTTGTATTCT TATTTCAGTG TTTCCTTCCA GACATTTTTA AGGTAATTGG CTTTAAAAAG   1774
```

AGAACATATT TTAACAAAGT TTGTG ACAC TCTAAAAAAT AAAATTGCTT TGTACTAGT 1833

SEQ ID NO: 4

SEQUENCE LENGTH: 508

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Met Pro Pro Lys Arg Asn Glu Lys Tyr Lys Leu Pro Ile
1 5 10

Pro Phe Pro Glu Gly Lys Val Leu Asp Asp Met Glu Gly Asn Gln Trp
15 20 25

Val Leu Gly Lys Lys Ile Gly Ser Gly Gly Phe Gly Leu Ile Tyr Leu
30 35 40 45

Ala Phe Pro Thr Asn Lys Pro Glu Lys Asp Ala Arg His Val Val Lys
50 55 60

Val Glu Tyr Gln Glu Asn Gly Pro Leu Phe Ser Glu Leu Lys Phe Tyr
65 70 75

Gln Arg Val Ala Lys Lys Asp Cys Ile Lys Lys Trp Ile Glu Arg Lys
80 85 90

Gln Leu Asp Tyr Leu Gly Ile Pro Leu Phe Tyr Gly Ser Gly Leu Thr
95 100 105

Glu Phe Lys Gly Arg Ser Tyr Arg Phe Met Val Met Glu Arg Leu Gly
110 115 120 125

Ile Asp Leu Gln Lys Ile Ser Gly Gln Asn Gly Thr Phe Lys Lys Ser

130 135 140

Thr Val Leu Gln Leu Gly Ile Arg Met Leu Asp Val Leu Glu Tyr Ile

145 150 155

His Glu Asn Glu Tyr Val His Gly Asp Val Lys Ala Ala Asn Leu Leu

160 165 170

Leu Gly Tyr Lys Asn Pro Asp Gln Val Tyr Leu Ala Asp Tyr Gly Leu

175 180 185

Ser Tyr Arg Tyr Cys Pro Asn Gly Asn His Lys Gln Tyr Gln Glu Asn

190 195 200 205

Pro Arg Lys Gly His Asn Gly Thr Ile Glu Phe Thr Ser Leu Asp Ala

210 215 220

His Lys Gly Val Ala Leu Ser Arg Arg Ser Asp Val Glu Ile Leu Gly

225 230 235

Tyr Cys Met Leu Arg Trp Leu Cys Gly Lys Leu Pro Trp Glu Gln Asn

240 245 250

Leu Lys Asp Pro Val Ala Val Gln Thr Ala Lys Thr Asn Leu Leu Asp

255 260 265

Glu Leu Pro Gln Ser Val Leu Lys Trp Ala Pro Ser Gly Ser Ser Cys

270 275 280 285

Cys Glu Ile Ala Gln Phe Leu Val Cys Ala His Ser Leu Ala Tyr Asp

290 295 300

Glu Lys Pro Asn Tyr Gln Ala Leu Lys Lys Ile Leu Asn Pro His Gly

305 310 315

Ile Pro Leu Gly Pro Leu Asp Phe Ser Thr Lys Gly Gln Ser Ile Asn

320 325 330

Val His Thr Pro Asn Ser Gln Lys Val Asp Ser Gln Lys Ala Ala Thr
335 340 345

Lys Gln Val Asn Lys Ala His Asn Arg Leu Ile Glu Lys Lys Val His
350 355 360 365

Ser Glu Arg Ser Ala Glu Ser Cys Ala Thr Trp Lys Val Gln Lys Glu
370 375 380

Glu Lys Leu Ile Gly Leu Met Asn Asn Glu Ala Ala Gln Glu Ser Thr
385 390 395

Arg Arg Arg Gln Lys Tyr Gln Glu Ser Gln Glu Pro Leu Asn Glu Val
400 405 410

Asn Ser Phe Pro Gln Lys Ile Ser Tyr Thr Gln Phe Pro Asn Ser Phe
415 420 425

Tyr Glu Pro His Gln Asp Phe Thr Ser Pro Asp Ile Phe Lys Lys Ser
430 435 440 445

Arg Ser Pro Ser Trp Tyr Lys Tyr Thr Ser Thr Val Ser Thr Gly Ile
450 455 460

Thr Asp Leu Glu Ser Ser Thr Gly Leu Trp Pro Thr Ile Ser Gln Phe
465 470 475

Thr Leu Ser Glu Glu Thr Asn Ala Asp Val Tyr Tyr Tyr Arg Ile Ile
480 485 490

Ile Pro Val Leu Leu Met Leu Val Phe Leu Ala Leu Phe Phe Leu
495 500 505

SEQ ID NO: 5

SEQUENCE LENGTH: 22

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

CTAATACGAC TCACTATAGG GC 22

SEQ ID NO: 6

SEQUENCE LENGTH: 21

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

TGTAGCGTGA AGACGACAGA A 21

SEQ ID NO: 7

SEQUENCE LENGTH: 21

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

TCGAGCGGCC GCCCGGGCAGG T 21

SEQ ID NO: 8

SEQUENCE LENGTH: 21

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

AGGGCGTGGT GCGGAGGGCGG T 21

SEQ ID NO: 9

SEQUENCE LENGTH: 22

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

CCAGGGTTTT CCCAGTCACG AC 22

SEQ ID NO: 10

SEQUENCE LENGTH: 22

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

TCACACAGGA AACAGCTATG AC 22

SEQ ID NO: 11

SEQUENCE LENGTH: 24

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

TGTAGTTCAG AAGAAGATTT GAGG 24

SEQ ID NO: 12

SEQUENCE LENGTH: 24

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

ATAATGGATC GCTTTGGGAG TGAC 24

SEQ ID NO: 13

SEQUENCE LENGTH: 26

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: ther nucl ic acid, synthetic DNA

SEQUENCE DESCRIPTION:

TGAAGGTCGG AGTCAACGGA TTTGGT 26

SEQ ID NO: 14

SEQUENCE LENGTH: 24

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

CATGTGGGCC ATGAGGTCCA CCAC 24

SEQ ID NO: 15

SEQUENCE LENGTH: 27

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

CCATCCTAAT ACGACTCACT ATAGGGC 27

SEQ ID NO: 16

SEQUENCE LENGTH: 24

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

GGATTTTCCT GATACTGTTT GTGG 24

SEQ ID NO: 17

SEQUENCE LENGTH: 24

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

ACCACAAACA GTATCAGGAA AATC 24

SEQ ID NO: 18

SEQUENCE LENGTH: 24

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

ACCTTTAAAA AGTCAACTGT CCTC 24

SEQ ID NO: 19

SEQUENCE LENGTH: 24

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

AAAAATTATC AAAAGGAATT TTGG 24

SEQ ID NO: 20

SEQUENCE LENGTH: 25

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

TTACTCTTAG TGAAGAGACA AACGC 25

SEQ ID NO: 21

SEQUENCE LENGTH: 36

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

AGCTGCGGCC GCGGTCTGCG GCTTAGGTGA AAATGC 36

SEQ ID NO: 22

SEQUENCE LENGTH: 36

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

AGCTGCGGCC GCAAAACAAA GAAAAGGAAA TCTGGT 36

SEQ ID NO: 23

SEQUENCE LENGTH: 37

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

AGCTGCGGCC GCAAGTGATG CCACCAAAAA GAAATGA 37

SEQ ID NO: 24

SEQUENCE LENGTH: 36

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

AGCTGCGGCC GCTGGAAGGA AACACTGAAA TAAGAA 36

SEQ ID NO: 25:

SEQUENCE LENGTH: 10

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu

1 5 10

SEQ ID NO: 26

SEQUENCE LENGTH: 33

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

GATCGATCCC GTCTGCGGCT TAGGTGAAAA TGC 33

SEQ ID NO: 26

SEQUENCE LENGTH: 60

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

GATGGATCCT TAGAGGTCTT CTTCTGAGAT GAGCTTCTGC TCCTTCTGGA CTCTCTTTCT 60

SEQ ID NO: 28

SEQUENCE LENGTH: 33

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

GATGGATCCA GTGATGCCAC CAAAAAGAAA TGA 33

SEQ ID NO: 29

SEQUENCE LENGTH: 60

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

GATGGATCCT TAGAGGTCTT CTTCTGAGAT GAGCTTCTGC TCGAGAAAAA ATAAAGCAAG 60

SEQ ID NO: 30

SEQUENCE LENGTH: 31

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

GATGGATCCC CATGCCTCGT GTAAAAGCAG C 31

SEQ ID NO: 31

SEQUENCE LENGTH: 31

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

GATGGATCCC CCAAAGAAAA GGAAATCTGG T 31

SEQ ID NO: 32

SEQUENCE LENGTH: 31

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

GATGGATCCC CATGCCACCA AAAAGAAATG A 31

SEQ ID NO: 33

SEQUENCE LENGTH: 31

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

GATGGATCCC CACAACATTT CGGTGAAGAG C 31

SEQ ID NO: 34

SEQUENCE LENGTH: 29

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: aingle

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

CCTTGTGTTG TATGGGTGGA ACCCAGTGA 29

SEQ ID NO: 35

SEQUENCE LENGTH: 31

SEQUENCE TYPE: nucleic acid

STRANDEDNESSS: single

TOPOLOGY: linear

MOLECULE TYPE: other nucleic acid, synthetic DNA

SEQUENCE DESCRIPTION:

ACACCACGAT GCCTGGAGCA ATGGCAACAA C 31

SEQ ID NO: 36:

SEQUENCE LENGTH: 25

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Met Leu Pro Glu Ser Glu Asp Glu Glu Ser Tyr Asp Thr Glu Ser Glu Phe Thr

1 5 10 15

Glu Phe Thr Glu Asp Glu Leu

20 25

SEQ ID NO: 37:

SEQUENCE LENGTH: 19

SEQUENCE TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION:

Cys Gln Thr Glu Glu Ala Ile Gln Thr Arg Ser Arg Thr Arg Lys Arg Val Gln

1 5 10 15

Lys

Claims

1. A protein having the amino acid sequence of SEQ ID NO: 2, or a protein having the same amino acid sequence where one or more amino acids are added, deleted, or substituted and having serine-threonine kinase activity.

2. A protein having the amino acid sequence of SEQ ID NO: 4, or a protein having the same amino acid sequence where one or more amino acids are added, deleted, or substituted and having serine-threonine kinase activity.

3. A protein encoded by a DNA sequence that hybridizes with the DNA sequence of SEQ ID NO: 1 or its complementary sequence and having serine-threonine kinase activity.

4. A protein encoded by a DNA sequence that hybridizes with the DNA sequence of SEQ ID NO: 3 or its complementary sequence and having serine-threonine kinase activity.

5. A DNA encoding the protein of any one of claims 1 to 4.

6. A vector comprising the DNA of claim 5.

7. A transformant carrying the vector of claim 6.

8. A method of producing the protein of any one of claims 1 to 4, wherein the method comprises cultivating the transformant of claim 7.

9. An antibody binding to the protein of any one of claims 1 to 4.

10. An antisense DNA against the DNA of claim 5 or part of it.

11. A method of screening compounds having inhibitory activity of serine-threonine kinase activity of the protein of any one of claims 1 to 4, wherein the method comprises

(a) contacting the protein of any one of claims 1 to 4 with a substrate to be phosphorylated by this protein in the presence of a test compound to detect the kinase activity of the protein of any one of claims 1 to 4, and
(b) comparing the kinase activity detected in step (a) with that detected in the absence of the test compound and selecting a compound that lowers the kinase activity of the protein of any one of claims 1 to 4.

Fig. 1

Adapter 1 5' CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT 3'

3' GGCCCGTCCA 5'

Adapter 2 5' TGTAGCGTGAAGACGACAGAAAGGGCGTGGTGCGGAGGGCGGT 3'

3' GCCTCCCGCCA 5'

Fig. 2

A.

Consensus sequence of the active site of serine-threonine kinase

[Leu,Ile,Val,Met,Phe,Tyr,Cys]-Xaa-[His,Tyr]-Xaa-Asp-[Leu,Ile,Val, Met,Phe,Tyr]-Lys-Xaa-Xaa-Asn-[Leu,Ile,Val,Met,Phe,Tyr,Cys]-[Leu, Ile,Val,Met,Phe,Tyr,Cys]-[Leu,Ile,Val,Met,Phe,Tyr,Cys]

B.

Consensus sequence of the ATP binding site of protein kinase

[Leu,Ile,Val]-Gly-Xaa-Gly-Xaa-[Phe,Tyr,Trp,Met,Gly,Ser,Thr,Asn,His] -[Ser,Gly,Ala]-Xaa-[Leu,Ile,Val,Cys,Ala,Thr]-Xaa-Xaa-[Gly,Ser,Thr, Ala,Cys,Leu,Ile,Val,Met,Phe,Tyr]-Xaa(5 times or 18 times)-[Leu,Ile, Val,Met,Phe,Tyr,Trp,Cys,Ser,Thr,Ala,Arg]-[Ala,Ile,Val,Pro]-[Leu,Ile, Val,Met,Phe,Ala,Gly,Cys,Lys,Arg]-Lys

Fig. 3

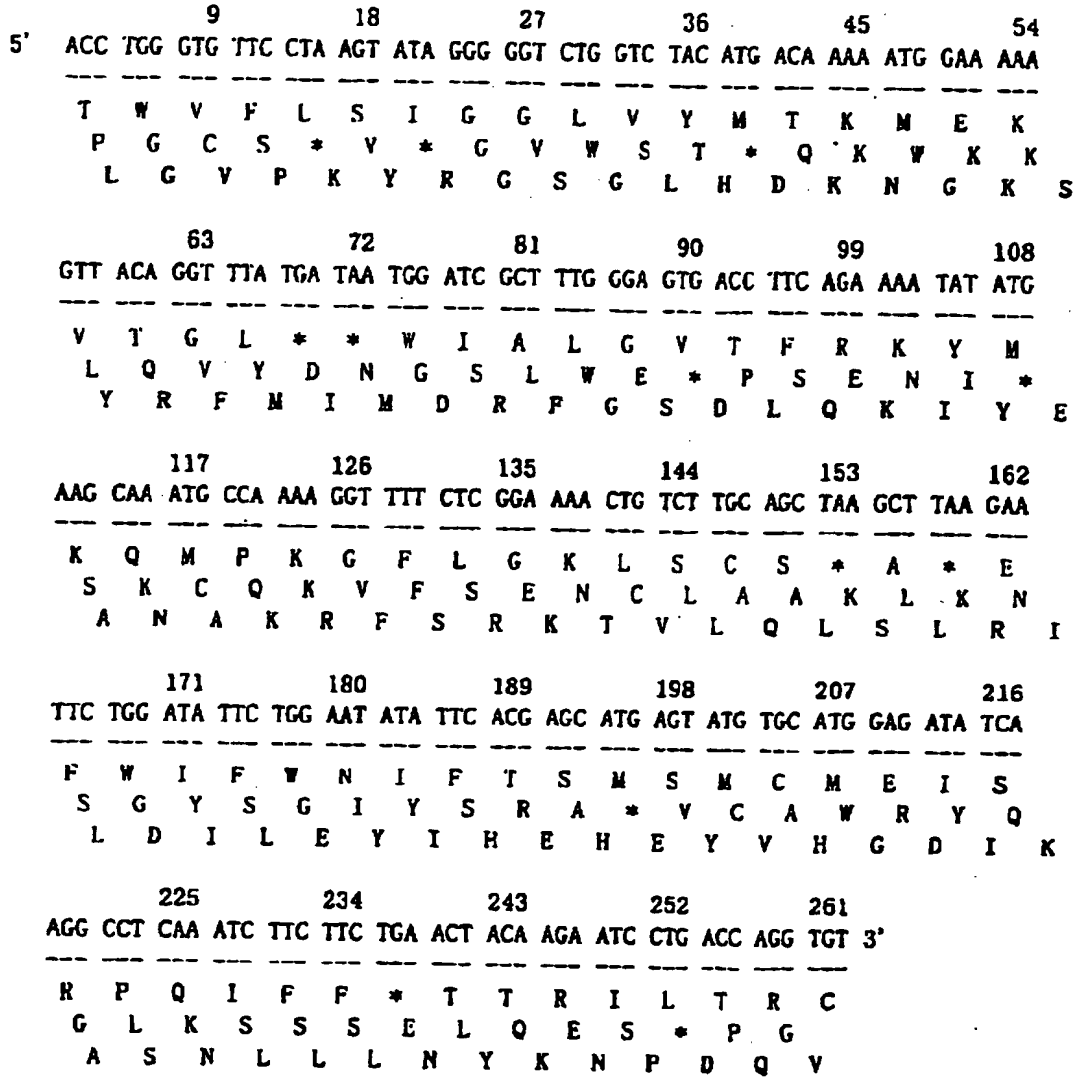


Fig. 4

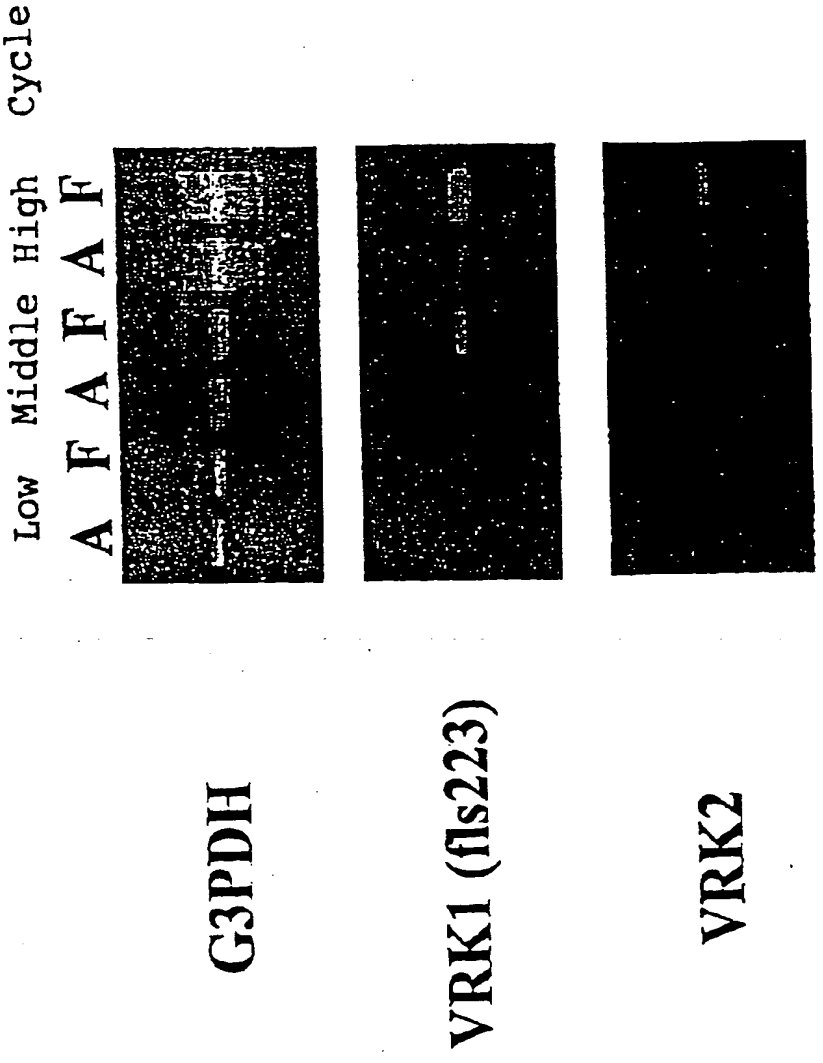


Fig. 5

```
                    10         20         30         40         50
VRK1       1 MPRVKAAQAG RQSSAKRHLA EQEAVGEIIT DMAKKEWKVG LPIGQGGFGC  50
B1R        1 M--------- ---------- -NFQ-GLVLT DNCKNQWVVG PLIGKGGFGS  50

                    60         70         80         90        100
VRK1      51 IYLADMNSSE SVGSDAPCVV KVEPSDNGPL FTELKFYQRA AKPEQIQKWI 100
B1R       51 IY-------- -TTNDNNYVV KIEPKANGSL FTEQAFYTRV LKPSVIEEWK 100

                   110        120        130        140        150
VRK1     101 RTRKLKYLGV PKYWGSGLHD KNGKSYRFMI MDRFGSDLQK IYEANAKRFS 150
B1R      101 KSHNIKHVGL ITCKAFGLYK SINVEYRFLV INRLGADLDA VIRANNNRLP 150

                   160        170        180        190        200
VRK1     151 RKTVLQLSLR ILDILEYIHE HEYVHGDIKA SNLLLNYKNP DQVYLVDYGL 200
B1R      151 KRSVMLIGIE ILNTIQFMHE QGYSHGDIKA SNIVIDQIDK NKLYLVDYGL 200

                   210        220        230        240        250
VRK1     201 AYRYCPEGVH KEYKEDPKRC HDGTIEFTST DAHNGVAPSR RGDLEILGYC 250
B1R      201 VSKFMSNGEH VPFIRNPNKM DNGILEFTPL DSHKGYVVSR RGDLETLGYC 250

                   260        270        280        290        300
VRK1     251 MIQWLTGHLP WED--NLKDP KYVRDSKIRV RENIASLMDK CFPEKNKPGE 300
B1R      251 MIRWLGGILP WTKISETKNC ALVSATKQKV VNNTATLLMT SL--QYAPRP 300

                   310        320        330        340        350
VRK1     301 IAKYMETVKL LDYTEKPLYE NLRDILLQGL KAIGSKDDGK LDLSVVENGG 350
B1R      301 LLQYITMVNS LTYFEEPNYD EFSHILMQG- ---------- ---------- 350

                   360        370        380        390        400
VRK1     351 LKAKTITKKR KKEIEESKEP GVEDTEWSNT QTEEAIQTRS RTRKRVQK.. 400
B1R      351 ---------- ---------- ---------- ---------- -----VYY.. 400
```

Fig. 6

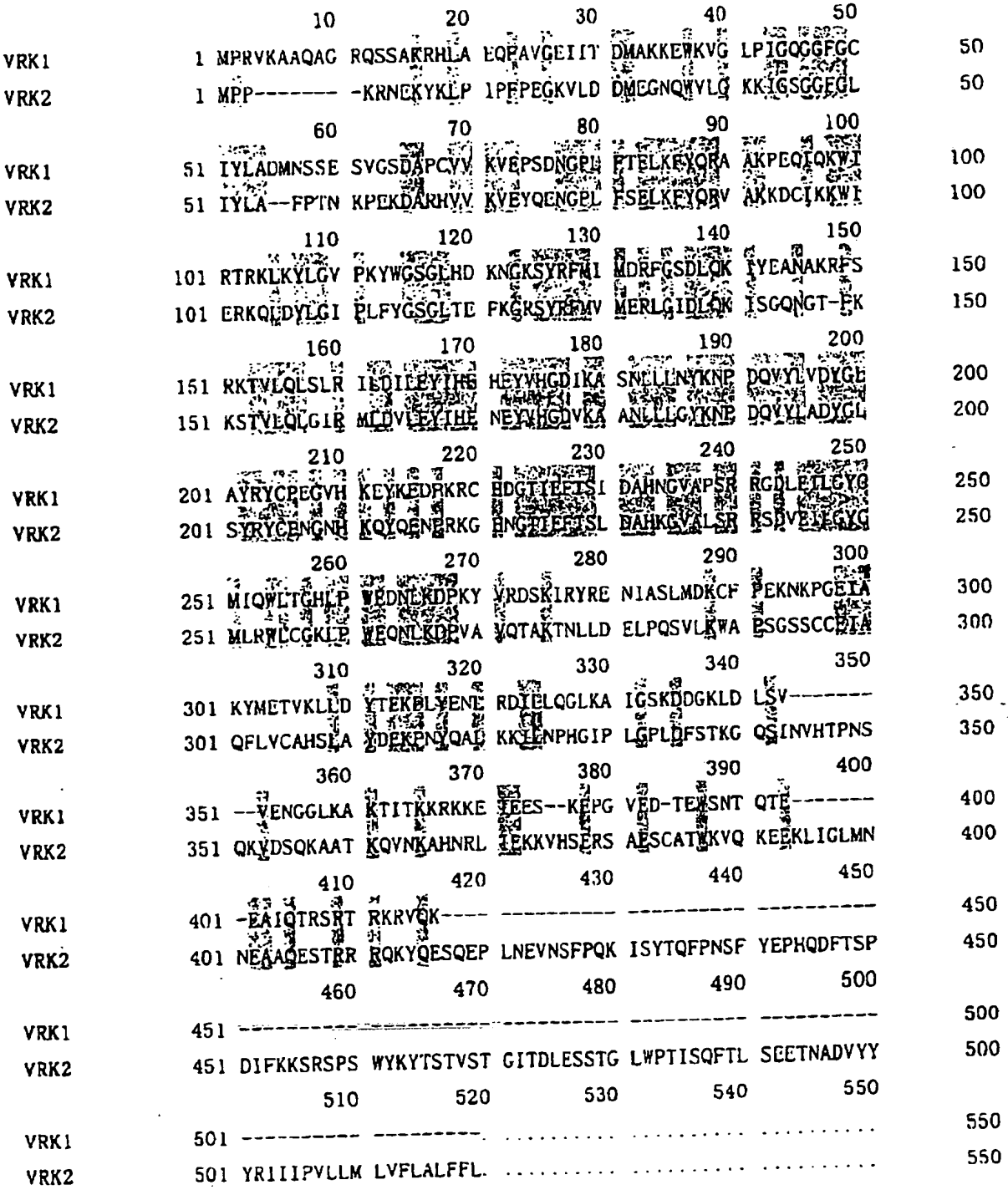


Fig. 7

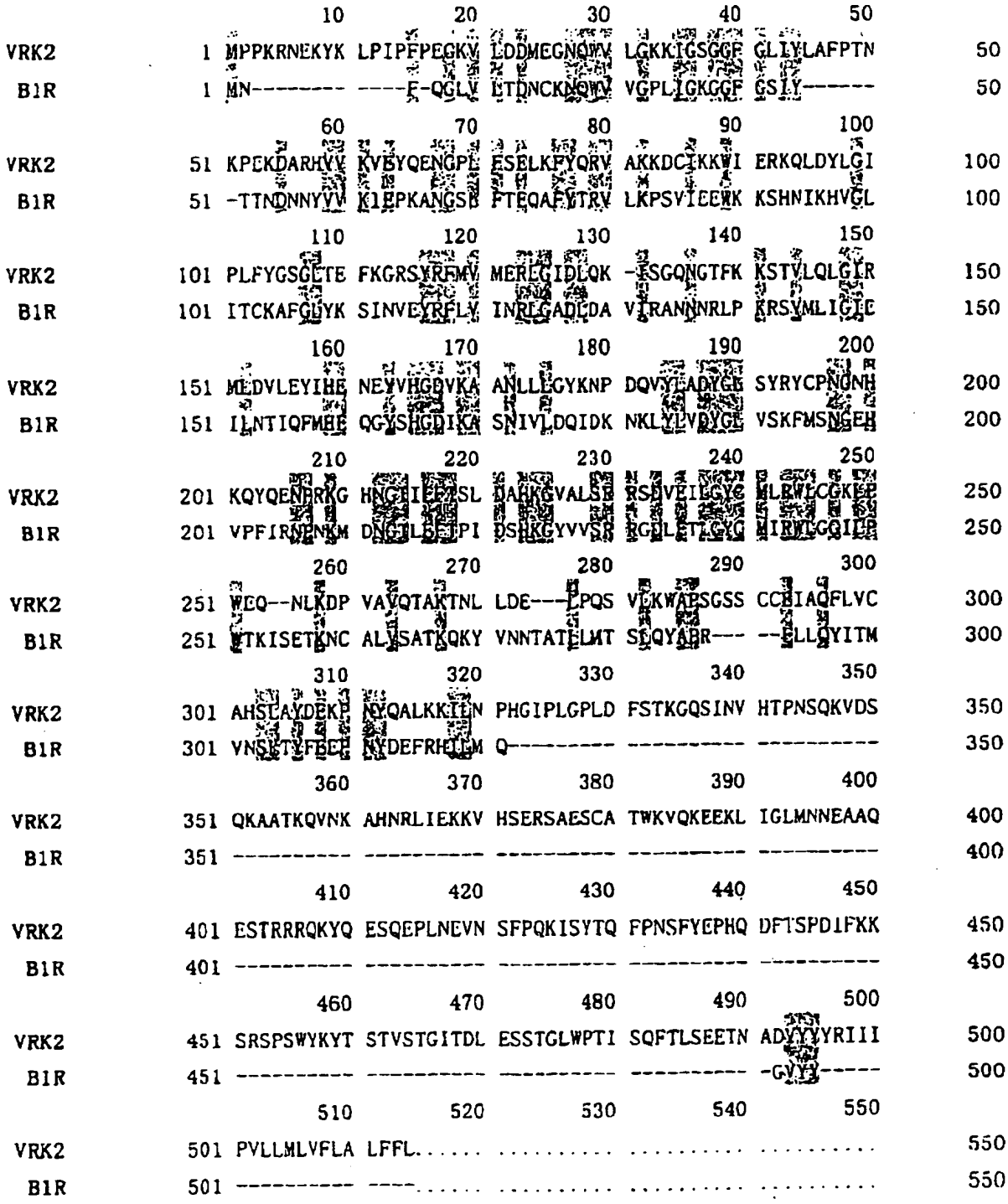


Fig. 8

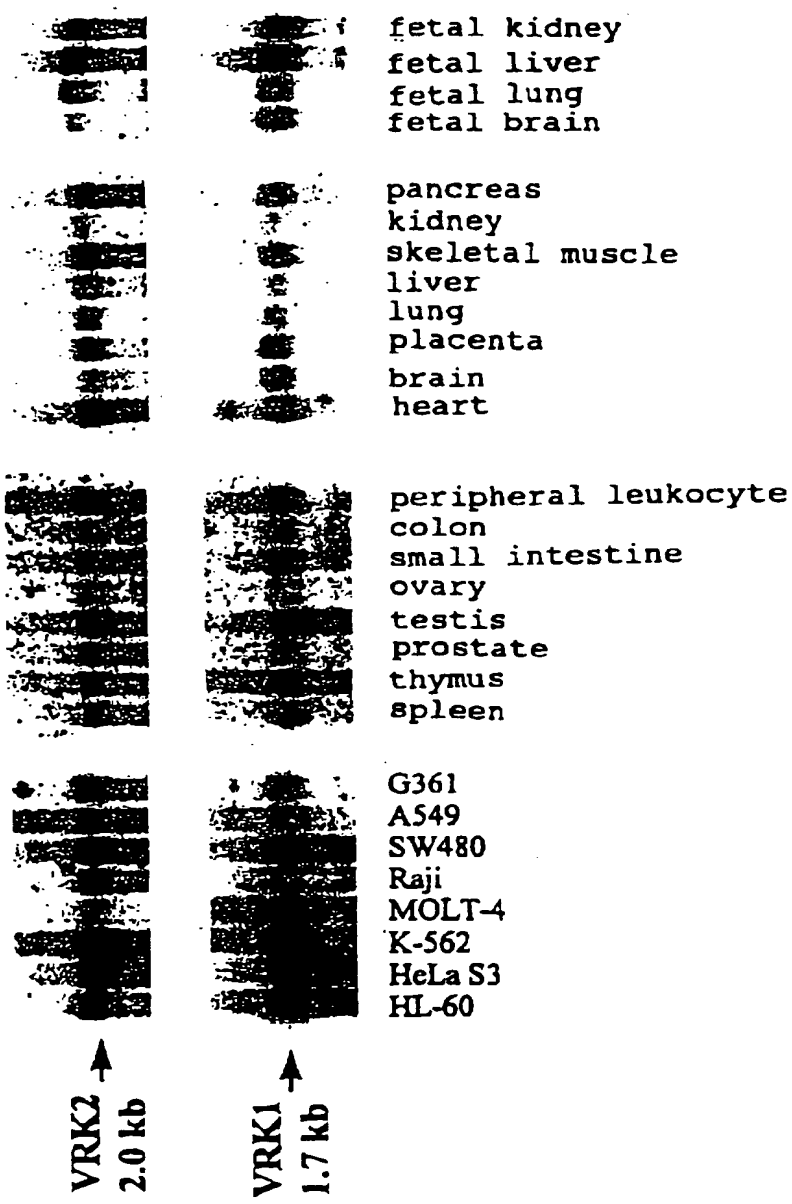


Fig. 9

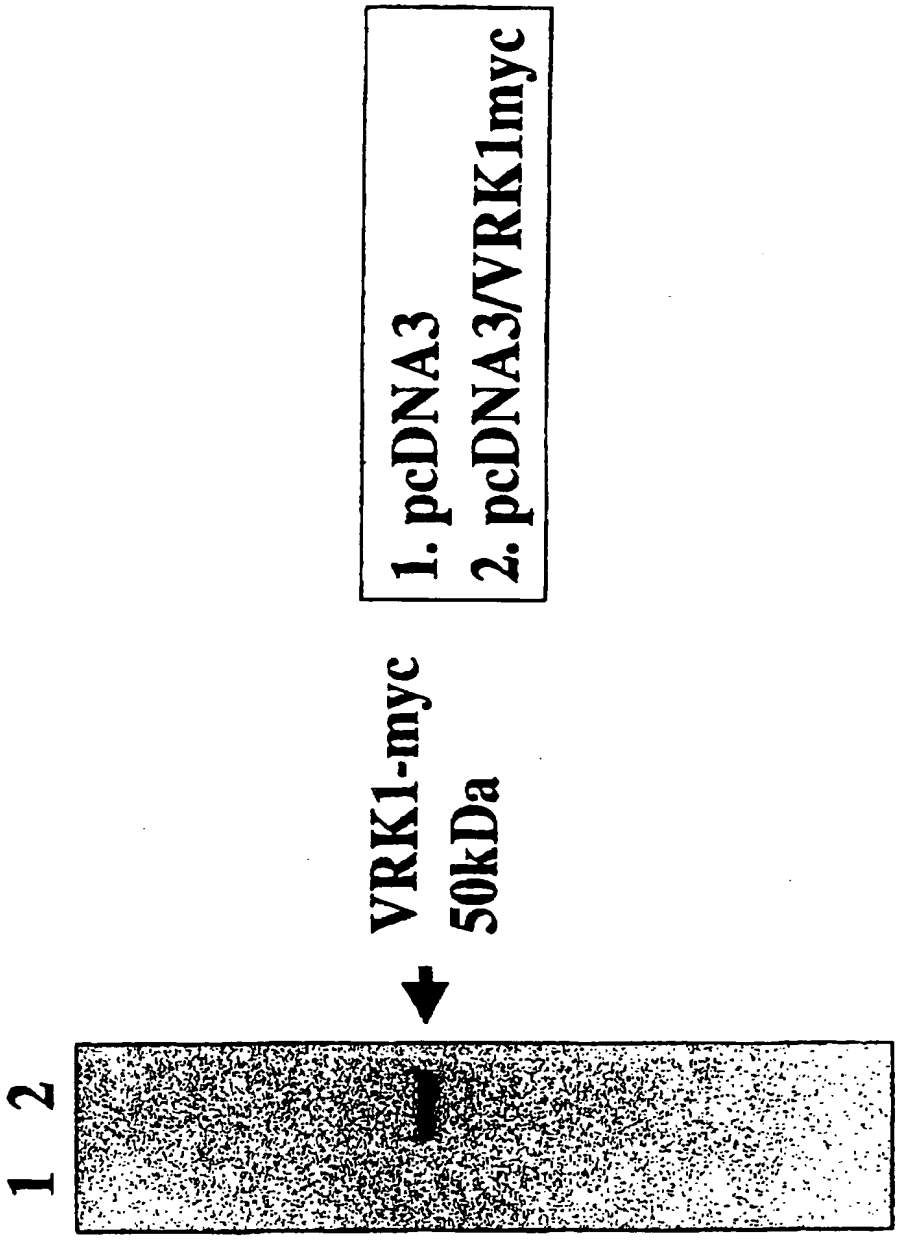


Fig. 10

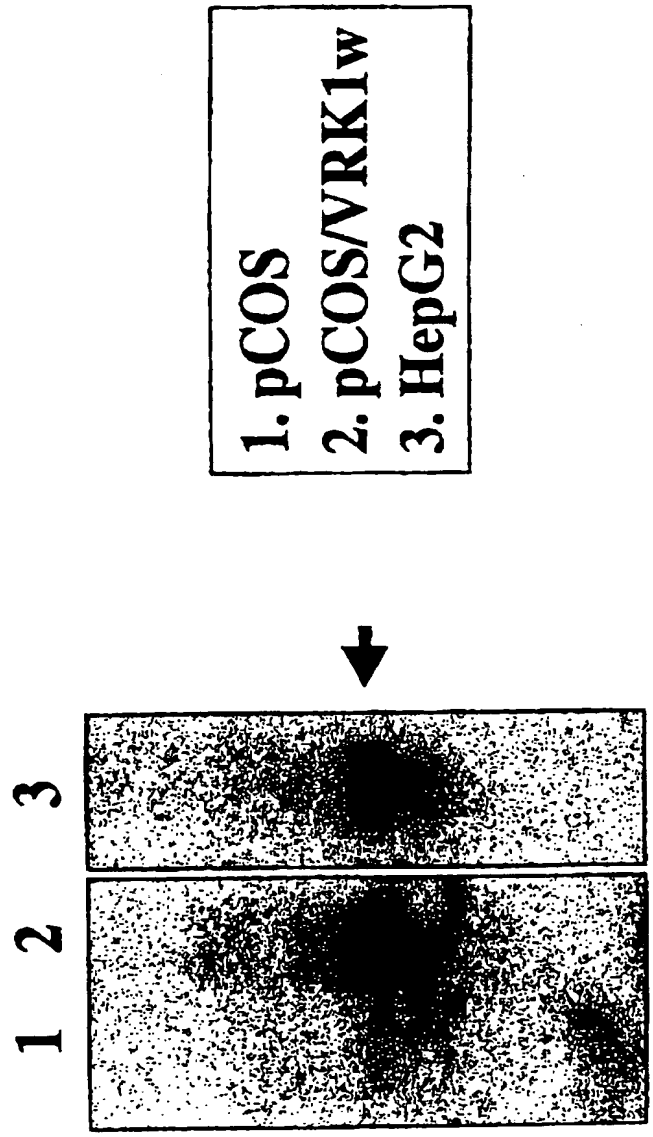


Fig. 11

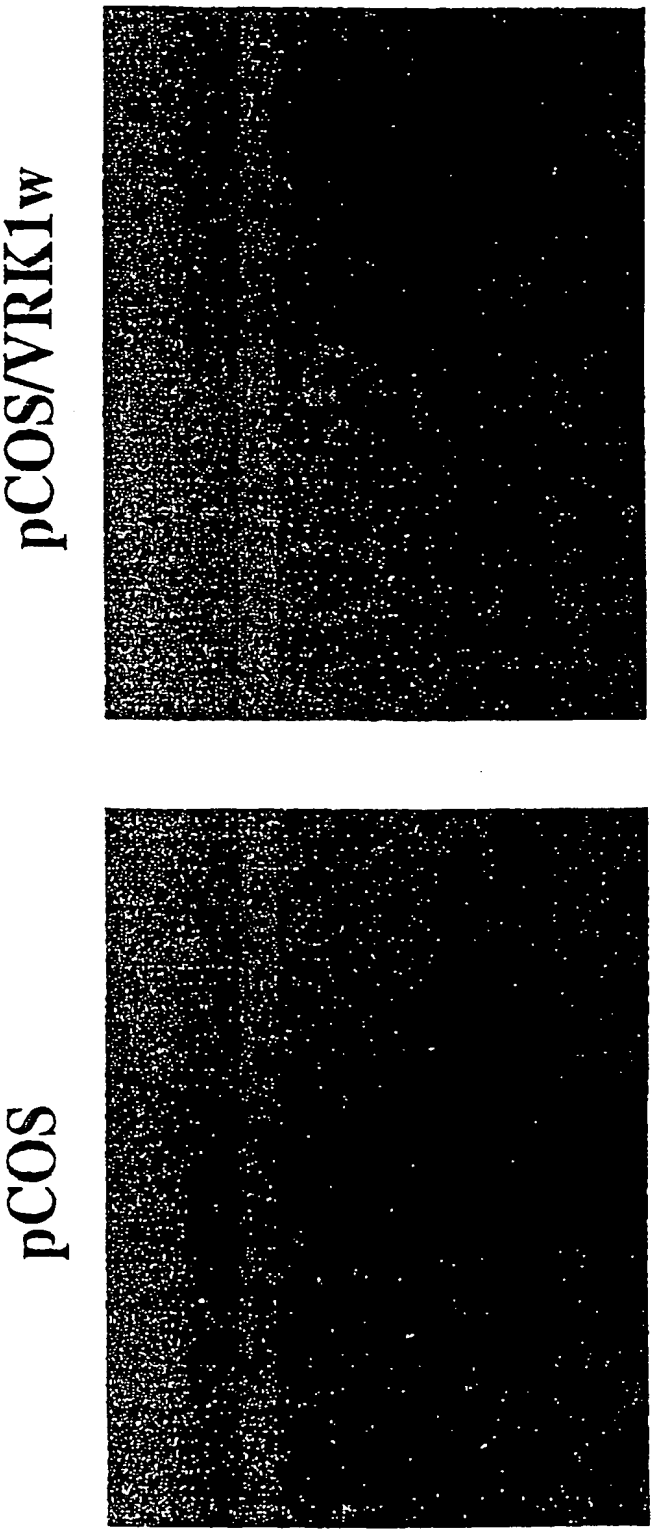


Fig. 12

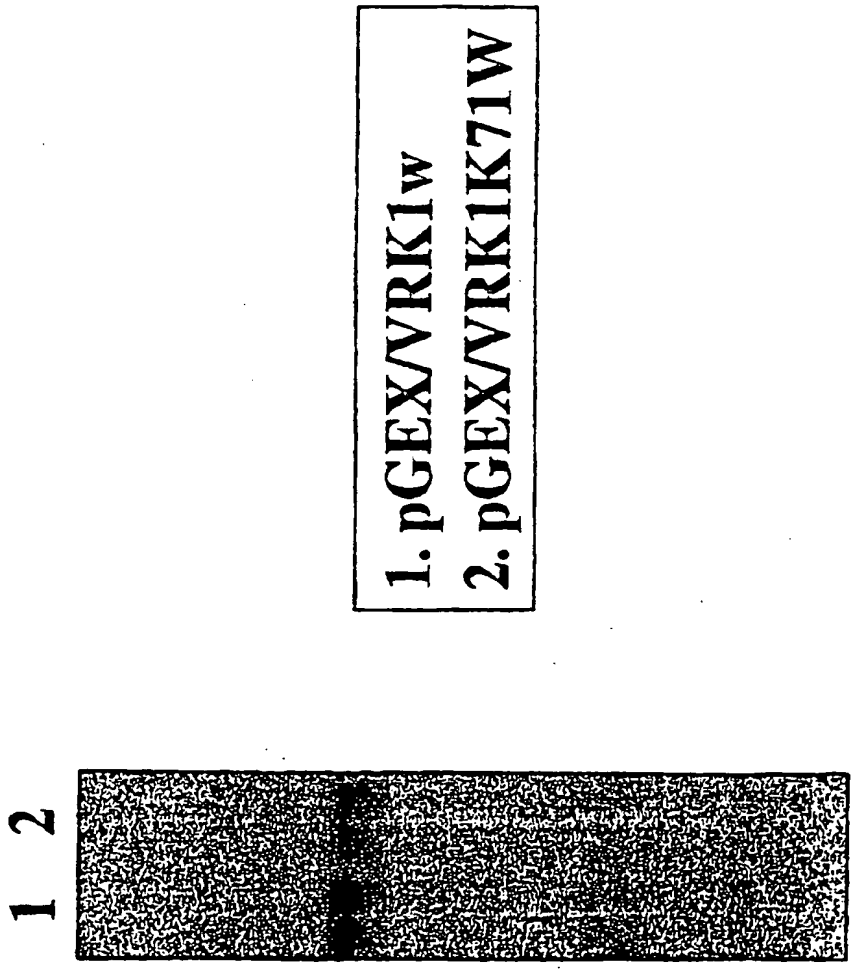


Fig. 13

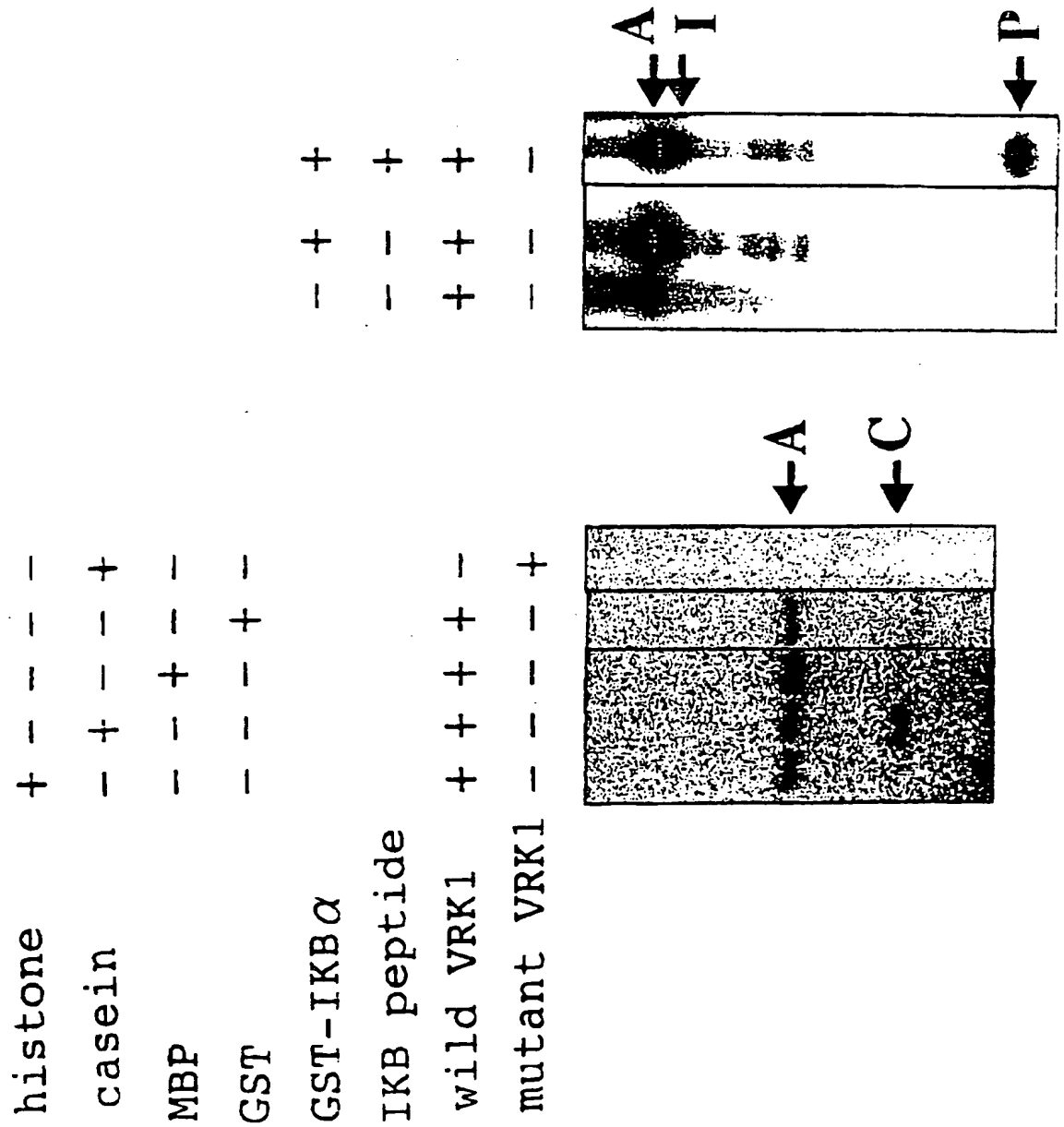


Fig. 14

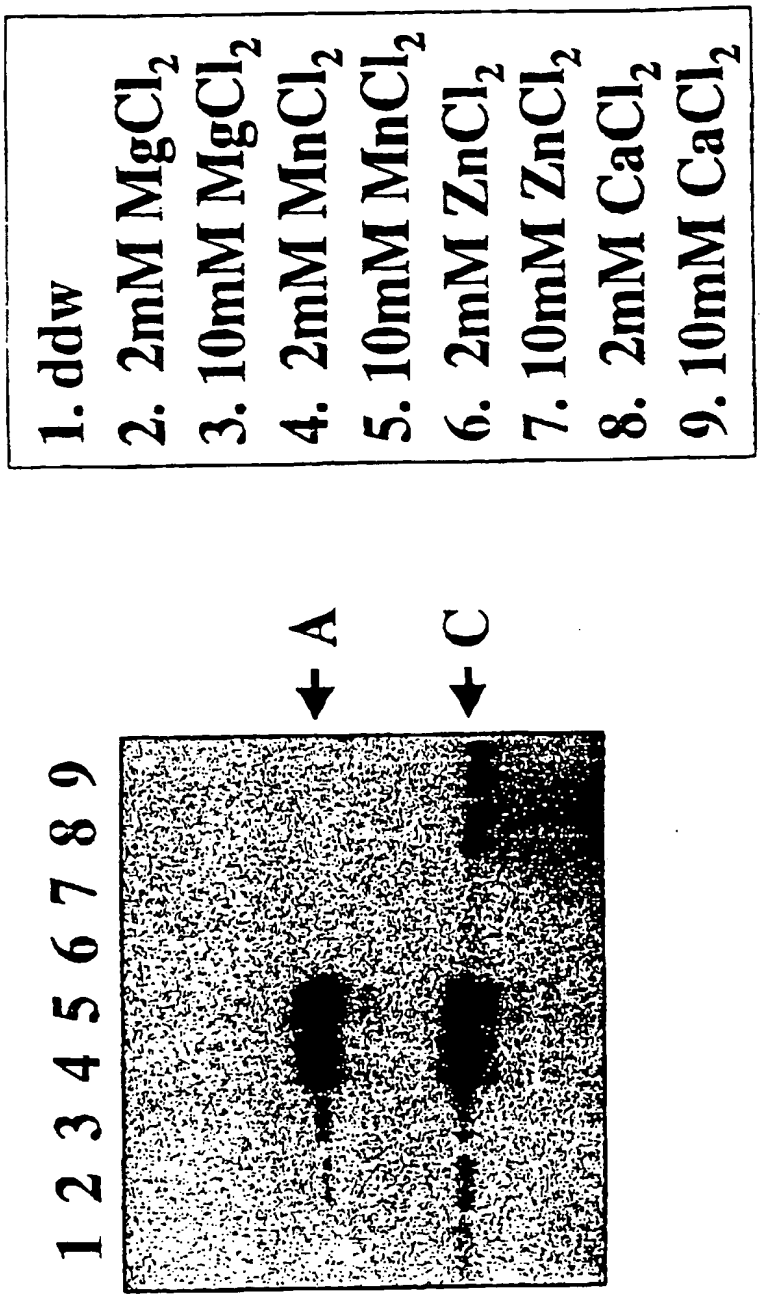
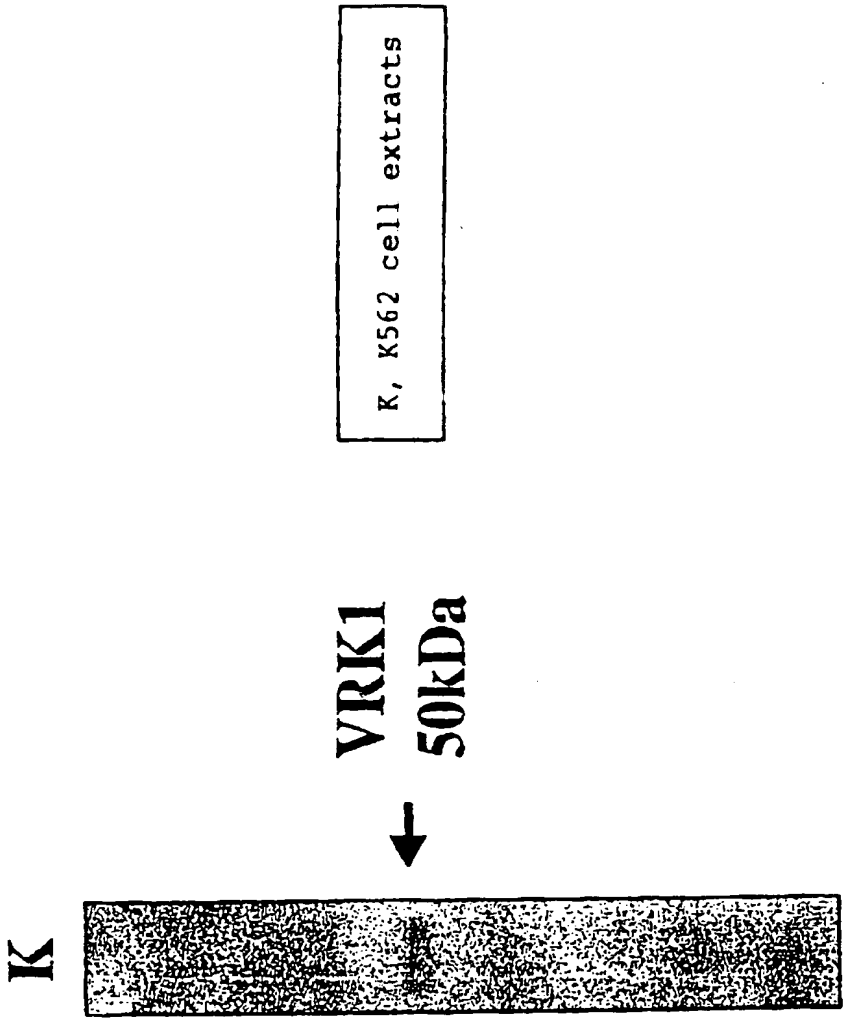


Fig. 15



INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP97/04855

A. CLASSIFICATION OF SUBJECT MATTER
Int. Cl^6 C12N15/54, C12N5/00, C12P21/02, C12Q1/48, C12N9/12

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int. Cl^6 C12N15/54, C12N5/00, C12P21/02, C12Q1/48, C12N9/12

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Genbank/EMBL/DDBJ(GNEETYX), BIOSYS(DIALOG), MEDLINE(STN)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	Genomics, Vol. 45(2)(1997-Oct) Nezu J-I et al. "Identification of two novel human putative serine-threonine kinases, VRK1 and VRK2, with structural similarity to vaccinia virus B1R kinase" p. 327-331	1 - 11
A	Proc. Natl. Acad. Sci. USA, Vol. 92(18)(1995) Lee N.H. et al. "Comparative expressed-sequence-tag ananlysis of differential gene expression profiles in PC-12 cells before and after nerve growth factor treatment" p. 8303-8307	1 - 11

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
February 2, 1998 (02. 02. 98)	February 10, 1998 (10. 02. 98)
Name and mailing address of the ISA/ Japanes Patent Office	Authorized officer
Facsimile No.	Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)